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

COVID-19 European primary care surveillance: Phased surveillance protocol

Author(s)	**Nivel and Epiconcept**
WP number	**2**
Deliverable number	**2.3**
Title of the deliverable	**Phased surveillance protocol**
Date	**15 June 2020**
DOI	**10.5281/zenodo.4569474**

Based on: current literature, I-MOVE primary care generic influenza vaccine effectiveness protocol 2019–2020

Table of contents

Version history

Version	Date finalised	Created/modified by	Comments
1.0	2020-04-24	Nivel/Epiconcept	Initial draft sent to partners
2.0	2020-06-01	Nivel/Epiconcept	Next version incorporating partners' comments

Abbreviations

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

➢ *The arrow indicates the sections that Member States 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). As of the 7th of April 2020, there were over 1.3 million cases of COVID-19 reported globally, with around half (608500) from the European Union/European Economic Area (EU/EEA) countries and the UK, and over 50000 of them have died. As of the 8th of April, the European Centre for Disease Prevention and Control (ECDC) reports that there is no indication at EU/EEA level that the peak of the epidemic has been reached.[1]

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 ten countries across the WHO European Region[1].

The WP2 primary care surveillance for COVID-19 is coordinated by Nivel (Netherlands institute for health services research). The I-MOVE-COVID-19 primary care network comprises nine sentinel surveillance networks in six European Union (EU) Member States (MS)[2] and in England and Scotland. The laboratory component of the network includes regional and national reference centres from the participating countries. While each of the surveillance sites can analyse their data separately, pooling the data for overall analysis will provide a sample size big enough to answer study questions with reasonable precision.

The I-MOVE-COVID-19 surveillance in primary care aims to reinforce and complement as much as possible ECDC's TESSy surveillance. This document presents the core European protocol for the primary care-based surveillance component of I-MOVE-COVID-19 for 2020, outlining the agreed methods for

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

[2] France, Ireland, The Netherlands, Portugal, Spain (two sites: the Spanish national system and the Navarra regional system) and Sweden.

collecting COVID-19 and SARS-CoV2 data during this pandemic. The specificities of each site's COVID-19 data collection can be detailed in the individual site protocol annexes. Other protocols are being developed for the primary care-based risk factor study for COVID-19 and for COVID-19 in healthcare workers (HCWs).

At time of writing (April-May 2020) the COVID-19 pandemic has caused changes in primary care seeking guidance and practices in some countries. Different scenarii, not all mutually exclusive, include:

- Strong lock-down with patients advised to stay at home, testing not possible and mostly or only teleconsultations at primary care level
- Set up of specific COVID-19 clinics for patients with COVID-19-like symptoms
- Set up of specific COVID-19 clinics, alongside sentinel GPs
- Limited testing capacity, due to lack of PPE or lack of tests and lab materials
- Adaptation of influenza surveillance systems to COVID-19

The aim of I-MOVE-COVID-19 is to support participating countries to progressively implement a system similar to influenza surveillance (as recommended by ECDC surveillance strategy [3]) and I-MOVE influenza. Where possible, within participating I-MOVE-COVID-19 primary care networks, a sample of patients suspected for COVID-19 will be tested for SARS-CoV-2. For each suspected COVID-19 case and each patient tested, information on patient characteristics is collected, and if possible also on symptoms and potential risk factors for COVID-19.

In their latest strategy for the surveillance of COVID-19[3], ECDC requests member states to report detailed data on confirmed cases, or at least weekly aggregated data on all confirmed cases. Within the primary care sentinel surveillance, it concerns:

1. the weekly number of SARS-CoV-2 positive ILI/ARI cases and the number of ILI/ARI tested
2. the weekly number of ILI/ARI consultations (numerator) and the corresponding denominator

This generic protocol will be updated according to the final surveillance strategy in each of the participating countries, and depending on the identification of new groups at risk. In particular, it aims at being flexible to be updated in the future to include estimation of COVID-19 VE, at the time when a vaccine becomes available.

This protocol is written in a generic manner and country-specific details of each study will be outlined in the study annexes (Annex 6).

2. General outline

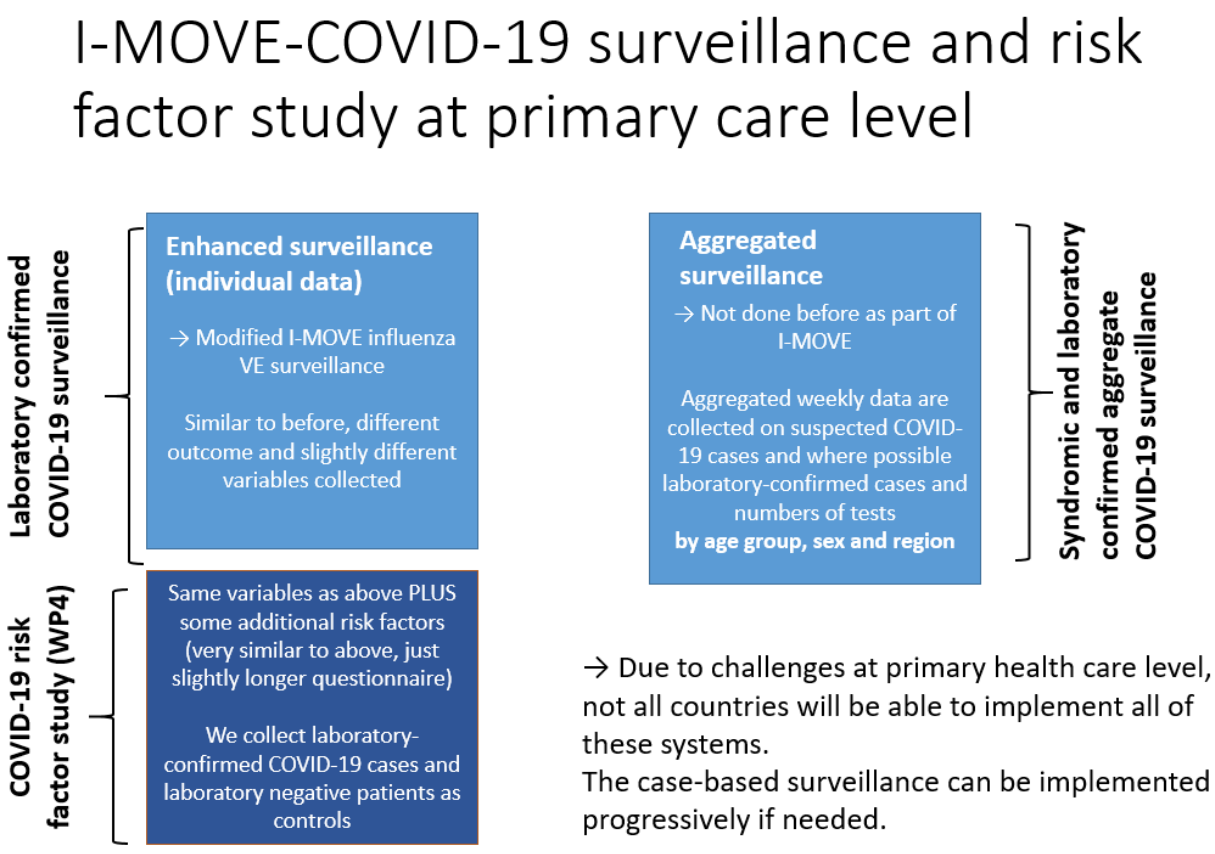
The I-MOVE network is nested in sentinel practices carrying out surveillance for influenza and now COVID-19. Sites part of the influenza I-MOVE network have practitioners used to carrying out enhanced surveillance to collect data for influenza vaccine effectiveness studies.

We will adapt the current I-MOVE influenza vaccine effectiveness protocol so that we can:

- Use it for enhanced COVID-19 surveillance
- Collect data for priority studies on risk factors for COVID-19 at primary care level
- Use the same protocol for influenza vaccine effectiveness (among sites taking part in I-MOVE influenza), and also COVID-19 vaccine effectiveness estimation once the vaccine becomes available

This I-MOVE-COVID-19 protocol is for enhanced COVID-19 surveillance, but much of the data collected within this surveillance are the same for the risk factors studies and the vaccine effectiveness studies, for which there are separate protocols. Those protocols are similar to this one, but collect more information. Where possible, we have indicated where further data for those studies will be collected.

Additionally, in order to meet the ECDC surveillance strategy on monitoring the intensity and geographic spread of COVID-19 in the population, sites can report weekly numbers of suspected COVID-19 cases, tests carried out and lab-confirmed COVID-19 cases. In order not to duplicate reporting to ECDC and to add value, reporting is carried out by age-group, sex and region. Where feasible, sentinel systems could be expanded to obtain a greater coverage for this surveillance, as recommended by the ECDC surveillance strategy for COVID-19.



Some countries/sites will be unable to carry out enhanced surveillance, due to various health care

related challenges. This protocol describes what is common for all levels of surveillance, what is specific for the enhanced surveillance and what is specific for the aggregated surveillance of weekly numbers of suspected and lab-confirmed COVID-19 cases. As the primary care systems change within the COVID-19 pandemic, sites can progressively implement the enhanced (case-based) surveillance.

3. Difference between enhanced (case-based) and aggregated surveillance

The enhanced (case-based) surveillance provides detailed information on a COVID-19 case and forms the basis of the risk factor study and the vaccine effectiveness study, once a vaccine is available. The enhanced surveillance is the core part of the primary care surveillance.

However some sites will not be able to collect the requested information for the enhanced (case-based) surveillance, due to changes in their healthcare system and new advice for healthcare seeking behaviour due to the pandemic. Or the complete information may not be available on a weekly basis. For these sites, the simpler aggregate reporting of tests carried out and numbers positive, or numbers of suspected COVID-19 cases, by age-group, sex and possibly region may be feasible. Note that if the extra aggregation by age-group and sex is not possible, then there is no difference to the ECDC TESSy COVID-19 aggregate reporting (NCOVAGGR), and data should be uploaded directly to TESSy. If the extra aggregation is possible, then the data will go to the I-MOVE-COVID-19 central hub, who will share the data from all participating sites with TESSy.

Some sites will be carrying out the enhanced (case-based) surveillance, but will only interview a sample of patients swabbed. In order to get a more complete picture, these sites can also use the aggregated surveillance to report on the totals of COVID-19 tests and positive cases.

Sites carrying out the enhanced (case-based) surveillance that interview all suspected COVID-19 patients swabbed do not necessarily need to participate in the aggregate surveillance. However there is added value in also participating in aggregated surveillance, as the frequency of reporting is weekly rather than monthly.

Participating countries/sites may also decide to expand the sentinel network for aggregated surveillance only.

4. Objectives

4.1. Primary objectives

The primary objective is to strengthen surveillance systems in nine study sites in seven European countries to rapidly detect and report COVID-19 cases in primary care, in order to contribute to the knowledge base, guide patient management, and inform the public health response for COVID-19 across Europe, by describing:

- Laboratory-confirmed COVID-19 cases by time, age group and region, and where possible by symptoms, patient characteristics and potential risk factors, and by virus (clades)
- Suspected COVID-19 cases by time, age group and region
- Proportion of confirmed cases by time, age group and region

To reach these objectives and depending on current study site situation, we propose the following data collection/reporting options:

Enhanced surveillance:

- Collection of individual level data on confirmed COVID-19 cases consulting at primary care level

Aggregated surveillance (by age, sex and region):

- Collection of aggregated weekly numbers of patients with suspected COVID-19, consulting in primary care (GP visit or telephone call)
- Collection of aggregated weekly numbers of SARS-CoV-2 tests carried out and aggregated weekly numbers of laboratory-confirmed COVID-19 at primary care level

➢ *Surveillance sites to define the objectives and types of their surveillance system*

4.2. Secondary objective of the enhanced (case-based) surveillance

The secondary objective of the enhanced (case-based) surveillance is to provide information to the I-MOVE-COVID-19 WP4 study on investigating risk factors for COVID-19 in primary care patients (Risk factors for COVID-19 at primary care level in Europe: Draft generic protocol), henceforth referred to as risk factor study.

4.3. Future and additional objectives

Additional future objectives for the primary care surveillance network include to potentially contribute to the following studies of I-MOVE-COVID-19 (WP4):

- Estimate the COVID-19 vaccine effectiveness (VE) in primary care patients of all ages (once a vaccine is available)[3]
- Investigate risk factors for COVID-19 in specific groups of patients at general practice level

5. General methods applicable to both enhanced and aggregated surveillance

5.1. Definition of surveillance of COVID-19 in primary care

Type of surveillance

- At surveillance site level: sentinel surveillance
- At European level: multicentre surveillance over several countries/regions

Population under surveillance

The surveillance population comprises community-dwelling individuals who consult a participating physician with symptoms of suspected COVID-19.

Surveillance period

The surveillance period starts in [month] 2020. Participating primary care practices carry out surveillance throughout the year.

Sample size

For COVID-19 primary care-based surveillance, there is no minimum sample size. Surveillance sites will conduct active surveillance to identify all or a systematic sample (depending on resources) of COVID-19 cases.

➢ *Surveillance sites to describe the setting (number of primary care practices included, number of primary care physicians, catchment population if possible). Note that this may be different for enhanced and aggregated surveillance.*

[3] This protocol will be updated to include estimates of VE once vaccine(s) become available.

- *Surveillance sites to define the beginning of primary care-based surveillance (date/month/year). If implemented already, please give the date of start.*

5.2. Outcomes

Enhanced case-based surveillance:

- The primary outcome of interest will be laboratory-confirmed COVID-19 in patients consulting at primary care level.
- The secondary outcome of interest is the genetic clade of SARS-CoV-2.

The outcomes for **aggregated surveillance** can include:

- Laboratory-confirmed COVID-19 in patients consulting at primary care level (aggregated weekly number)
- Suspected COVID-19 in patients consulting at primary care level (aggregated weekly number)

5.3. Case definitions

Patients are persons consulting a general practitioner, defined as someone either

- Having a face-to-face consultation with the practitioner (in the practice or at the patient's home)
- Having a telephone/video consultation with the practitioner[4]

A **suspected COVID-19 case** is defined as a patient with:

- acute onset of at least one of the following symptoms: fever or cough or sore throat or shortness of breath or coryza

or

- a clinician's judgment that illness might be due to a SARS-CoV-2 infection

A **confirmed COVID-19 case** will be defined as a suspected COVID-19 case with a respiratory sample positive for SARS-CoV-2.

A **COVID-19 negative case** will be defined as a suspected COVID-19 case with a respiratory sample negative for SARS-CoV-2. These suspected COVID-19 patients negative for SARS-CoV-2 will be used in I-MOVE-COVID-19 studies, including determining risk factors for COVID-19 and for vaccine effectiveness

[4] For enhanced surveillance, we can include these patients if a swab can be taken soon after the consultation (either by the patient self-swabbing, visiting a specific swabbing centre or the practitioner taking a swab, either at patient's home or at the general practitioner's office)

(once a vaccine is available).

- *Surveillance sites to document the case definitions they use.*

5.4. Exclusion criteria of cases for surveillance

The following patients will be excluded from surveillance if the surveillance site country requires consent and this was refused, e.g. if the patient or his or her legal guardian is unwilling to participate or unable to communicate and give consent according to the local ethical review process.

- *Surveillance sites to document reasons for exclusion of potential eligible patients.*

5.5. Case identification

Following the procedures outlined by each study, all suspected COVID-19 cases are selected and asked to provide a nasal/throat swab specimen for SARS-CoV-2 testing. Sampling all suspected COVID-19 cases is preferred, in particular all patients aged 65 and over, but if this is not possible, then a systematic sample can be taken, e.g. the first three suspected COVID-19 cases seen each week per GP stratified by age (0-64 years and 65 years or older). SARS-CoV-2-positive COVID-19 patients are considered as lab-confirmed COVID-19 cases.

- *Each surveillance site to describe the procedures to select suspected COVID-19 patients to swab.*

5.6. Laboratory methods

Primary care practitioners will collect respiratory specimens (see Section 5.5) from either all or a systematic sample of suspected COVID-19 patients consulting a practitioner and consenting to take part in the study, respecting safety standards for COVID-19 and following WHO biosafety guidelines.[5] Depending on the setting, some practitioners will refer patients to specific COVID-19 testing centres, or some patients may even be carrying out self-swabbing at home.

A comprehensive generic laboratory protocol will be developed and presented in the future alongside this surveillance protocol.

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

- *Each surveillance site to describe the type (nasopharyngeal/oropharyngeal or both) and number of swabs taken for each patient.*
- *Each surveillance site to describe where swabbing will be carried out (at practice, at home, in centres, a mixture).*

Quality control tests should systematically be run using PCR to ensure presence of cells in the respiratory specimens. In the absence of cells, a negative result should be considered inconclusive and a second swabbing should take place if possible.

The ECDC-recommended SARS-CoV-2 laboratory confirmation is by viral RNA detection with nucleic acid amplification tests, such as RT-PCR.(1,4) Isolates will undergo molecular analysis for currently circulating SARS-CoV-2 virus. During the influenza season, tests should also be performed for influenza viruses 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 random sample thereof. The sample should be random and thus representative of cases and be large enough to provide reasonable precision when calculating proportions of virus change over time. Gene sequences should also be uploaded to GISAID's open access EpiCoV platform. Gene sequence information can be provided directly to the I-MOVE-COVID-19 central hub, or the GISAID EpiCoV accession number can be provided alongside the I-MOVE-COVID-19 unique identifier to link these data (see annex 3). Processed genetic information, e.g. name of genetic clade, can also be included within the epidemiological database.

- *Each surveillance site to describe the laboratory procedures (samples taken, storage, transport).*
- *Each surveillance site to describe the tests (including point-of-care tests) and the kits used (and their sensitivity, specificity, PPV) for COVID-19 and, if needed, other respiratory virus detection.*
- *Each surveillance site to describe if the laboratory participates in QA/QC (Quality Assurance/Quality Control) schemes.*
- *Each surveillance site to describe the selection of specimens and the methods for genetic and, when it becomes available, antigenic characterisation.*
- *Each surveillance site to describe genetic and, when it becomes available, antigenic analyses and specify sequencing methods.*

6. Overview of data collection for enhanced surveillance

6.1. Minimum dataset for enhanced surveillance

The minimum collected information should include:

- unique identifiers for GP and individual patients
- demographics: age, sex, region
- healthcare worker (yes/no)
- COVID-19 signs, symptoms
- date of onset of symptoms
- referral by GP to hospital
- swab/test information: date of swabbing; laboratory results (including genetic, and when it becomes available, antigenic analysis, where available)
- selected underlying chronic conditions
- pregnant (yes/no)
- smoking history
- travel history in last 14 days
- current seasonal influenza vaccination
- pneumococcal vaccination status

More details on these variables are described in section 8.

See also Annex 1: List of variables, definition and coding for enhanced surveillance.

6.2. Additional information for enhanced surveillance

Additional information to be collected for surveillance could include:

- type of swab (nasal, throat)
- information on test (PCR, point of care)
- results for other respiratory pathogens (RSV, rhinovirus, human metapneumovirus, adenovirus, seasonal coronaviruses, bocavirus)
- information on other chronic diseases

6.3. Additional information for risk factor study

Additional information for the risk factor study (optional) could include:

- urban/rural environment
- number of hospitalisations for the chronic diseases in the previous 12 months
- number of GP visits in the previous 12 months
- current season influenza vaccination date and product

- type of pneumococcal vaccine and either date or year of vaccination
- BCG vaccination
- outcome: was the patient hospitalised or did the patient die, as part of this illness episode
- functional status
- pre-swabbing antiviral administration (including date and type)
- presymptomatic medication
- protective factors used
- contact to a confirmed case
- exposures in the past 14 days
- size of household
- previous Sars-CoV-2 test (including date and result)

See also section 8 and Annex 2.

➢ *Each study site to list the variables collected.*

6.4. Data collection instruments for enhanced surveillance

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

- face-to-face/telephone interview
- electronic medical records
- interview with patient or his/her family
- vaccination register
- laboratory

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

7. Overview of data collection for aggregated surveillance

7.1. Aggregated surveillance data on numbers of tests and patients testing positive

Study sites able to collect the number of laboratory-confirmed COVID-19 cases in primary care, to report on a weekly basis:

- numbers of tests carried out and

- number of patients tested positive

stratified by:

- age group (0-14 years, 15-44 years, 45-64 years, 65+ years)
- sex
- country or region

This is an extension of the TESSy NCOVAGGR upload of total number of patients tested and total number of patients tested positive.

➢ *Sites to define the types of consultations and the source population.*

➢ *Sites to define the regions used for surveillance.*

➢ *Sites to define any further stratification variables used (e.g. presence of chronic condition)*

7.2. Aggregated syndromic surveillance data on suspected COVID-19 cases

Additionally, or alternatively, GPs may be asked to report each week the total number of suspected COVID-19 cases consulting by the same stratification variables:

- age group (0-14 years, 15-44 years, 45-64 years, 65+ years)
- sex
- country or region

The denominator for each stratum could be the source population or the total number of consulting patients.

Further practice-level information is collected as well and updated whenever there is a change:

- Type of consultations that the majority of surveillance data is collected by: face-to-face, teleconsultations, help lines, home visits.
  If needed, the surveillance information can be collected by type of consultation.
- Source population of the surveillance.
- Source of data collection: aggregated reporting (by the GP) or extracted from electronic medical records.

➢ *Sites to define the types of syndromic data used and the source population.*

➢ *Sites to define the regions used for surveillance.*

➢ *Sites to define any further stratification variables used (e.g. presence of chronic condition)*

7.3. Expansion of sentinel network for aggregated surveillance

In line with ECDC's COVID-19 surveillance strategy, where feasible sentinel surveillance systems should be expanded to include more physicians to achieve greater population coverage. This may apply to the suspected COVID-19 surveillance only, or to both the suspected and lab-confirmed COVID-19 surveillance.

- *Sites to describe expansion of sentinel surveillance, if any.*

8. Detailed information on data collection

This part of the protocol describes details on the data collection. Most data is collected specifically for the enhanced surveillance at individual (case-based) level. Please see section 7 for the restricted list of variables for the aggregated surveillance.

As the risk factor study will be based on the enhanced surveillance, the extra information to be collected for the risk factor study is listed in section 8.2.

8.1. Data collection for enhanced surveillance

Patient characteristics

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

- Age in years
- Sex
- Region
- Smoking history (never smoked, former smoker (stopped smoking for at least one year), current smoker (including stopped smoking less than one year ago)
- Pregnancy (yes/no)
- Healthcare worker (yes/no)

Region refers to NUTS2 areas within a country and we propose to use the GP location as proxy. NUTS2 areas might be regions or provinces. See: https://ec.europa.eu/eurostat/web/nuts/background

Smoking refers to any type of smoking (cigarettes, cigars, vaping, etc.).

The definition of a healthcare worker for the purposes of this surveillance is a person who is working (paid or on a regularly voluntary basis) in healthcare AND has been in contact with patients (any type of

patient) during his/her work. This includes: doctors, nurses, emergency medical personnel, medical and nursing students with contact to patients, porters and cleaners.

Information on consultation

- Type of consultation: in practice, video, telephone, home
- Date of consultation

Clinical signs and symptoms

We will collect information on symptoms to better understand the clinical spectrum of disease. It is also important for the VE studies to collect symptoms as completely as possible.

As a minimum:

- date of symptoms onset
- fever/feverishness
- if fever: measured fever (temperature)
- cough
- shortness of breath
- anosmia
- ageusia
- dysgeusia
- headache
- sore throat
- fatigue
- myalgia
- malaise
- coryza, rhinitis
- chest pain
- chills

If possible, sites could also collect:

- nausea
- vomiting
- diarrhoea
- stomach ache
- conjunctivitis
- dizziness
- cyanosis or associated pulse oximetry
- rash or other dermatological manifestation
- palpitations

Information on swabbing and test results

For each patient we will collect information on:

- date of swabbing
- result of COVID-19 test

If possible, sites can collect information on:

- test results from any other respiratory viruses
- place of swabbing (GP practice, COVID-19 centre, self-swabbing)
- type of swab (nasal, throat, both)
- type of COVID-19 test (PCR, point-of-care)

Pre-existing chronic conditions

The list of underlying conditions should include at least:

- diabetes (optionally sites can distinguish between type 1 and type2)
- cardiovascular disease (myocardial infarction, angioplasty, coronary artery bypass surgery, stroke, transient ischemic attacks, treated hypercholesterolemia)
- chronic pulmonary disease (not including asthma)
- asthma
- cancer
- obesity
- immunodeficiency

If physicians are recruiting cases using electronic medical records, the list of codes (ICD or ICPC) can be used to document a study participant's chronic diseases (see Table 1).

If codes are not available, a list of underlying conditions should be prepared by using a short questionnaire.

For obesity, we will collect body mass index (BMI). If it is possible to collect the actual BMI or height and weight, this is preferred. If not possible, we suggest categories (BMI: 30–39 and ≥40).

Information on additional chronic conditions could be collected either as part of the surveillance or as part of the risk factor study:

- hypertension
- renal disease
- liver disease
- rheumatological diseases

Vaccination status

For the surveillance study, we will collect information on previous influenza and pneumococcal vaccination:

- seasonal influenza vaccination from the most recent influenza season (with date, if possible)
- latest pneumococcal vaccination type (with date or at least year, if possible)

The sources of information for vaccination may include:

- vaccination registry
- consultation of the patient's vaccination card
- GP electronic medical records
- self-report
- data from the patient's insurance company showing evidence of pharmacy delivery or reimbursement of influenza vaccine during the current influenza season.

Travel

To determine potential place of infection, we will ask about the patient's recent travel (this question will be more relevant after social distancing measures end):

- Have you travelled outside this country in the past 14 days?
- If yes, which country?

Referral to hospital

In order to get a measure of severity of the case, we will include the question:

- Did you recommend the patient be referred to hospital?

➢ *Each surveillance site to provide a complete list of information they can collect, including documentation on how the information is being collected.*

Table 1. ICD-9 and ICD-10 codes for chronic diseases [QUESTION TO STUDY SITES: ARE ICD-CODES USED? IF YES, COULD YOU PLEASE REVIEW?]

Category	ICD-9	ICD-10
Anaemia	280–285	D50-64
Asplenia	746.87, 759.0	Q89.01, Q20.6, Z90.81
Asthma	493.0, 493.1, 493.9	J45
Chronic liver disease	571	K70, K72-74, K754, K769

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
Diabetes	250	E10-11
Hypertension	401, 401.0, 401.9, 405, 405.91, 405.99,	I10, I15.8, I15, I15.1, I15.2, I97.3, I27.0
Obesity	27800, 278.01, 278.03	E66.01, E66.2, E66.9
Immunodeficiency* or organ transplant	042, 279, V08, V42	B20, D80-84, D89.8-9, Z21, Z94
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,
Renal disease	274.1, 408, 580–591, 593.71–593.73, 593.9	M10.30, N00-19, N20.0, N28.9
Dementia	290, 294, 331	F01, F03, F05, G30, G31, G91, G94
Stroke	348, 438	G93, I67.83, I69
Rheumatologic diseases	446, 710, 714	M30-34, M35.0, M35.5, M35.8-9, M05-06, M08, M12.00
Cancer	140–208	C00-96
Lung disease excluding asthma)	011, 490–511 (exclude asthma), 512.8, 513–517, 518.3, 518.8, 519.9, 714.81	A15, J40–44 J46–47, J60–94, J96, J99, J182, M34.81, M05.10
Tuberculosis		A15–A19

* Patients who are only treated with glucocorticoids and have no other immune deficiency, are considered immune suppressed when treated with high-dose corticosteroids (≥ 20 mg/day of prednisone or equivalent for ≥2 weeks) in the last 3 months.

8.2. Data collection for the risk factor study (optional)

The following information will only be collected for the risk factor study, but are included in this protocol for completeness.

Further patient characteristics

- Level of urbanisation of the place of residence
- Deprivation score

The level of urbanisation of the place of residence may be a risk factor for COVID-19. We propose to use GP practice location as a proxy, rather than patient location of residence.

We suggest collecting urban/rural information according to the Eurostat NUTS3 regions (https://ec.europa.eu/eurostat/web/rural-development/methodology) with three classifications: predominantly rural, intermediate, predominantly urban.

Study sites can collect this information in several ways:

- Asking the GP to provide the classification of urban/rural directly, upon specifications.
- Obtaining the GP practice postcode and then translating it to the appropriate urban/rural classification.

Deprivation may also be a risk factor for COVID-19. We propose to use GP practice location as a proxy, rather than patient location of residence. Where available, we suggest countries to calculate from the GP postcode the European deprivation index for easier standardisation across countries (https://academic.oup.com/eurpub/article/28/suppl_4/cky213.625/5191925).

Antiviral administration before swabbing

The use of antivirals prior to swabbing may lead to misclassification biases in studies where test-negative controls are used. We will document whether the patients received any antiviral treatment in the 2 weeks preceding symptom onset and the type (curative or preventive) of antivirals received.

Presymptomatic medication/vaccinations

- angiotensin-converting enzyme (ACE) inhibitors
- angiotensin II receptor blockers (ARBs)
- non-steroidal anti-inflammatory drugs (NSAIDs)
- statins
- corticosteroids
- biological disease-modifying anti-rheumatic drugs (DMARDs)
- current/recent chemotherapy
- metformin
- antithrombotic/ platelet aggregation inhibitors
- BCG vaccination

Information on previous positive COVID-19 infections

Among those patients consulting their GP with COVID-19-like symptoms, some may have already had a positive COVID-19 test in the past for a previous SARS-CoV-2 infection. This gives information on those who have had more than one SARS-CoV-2 infection and also helps with the control selection for the risk factor study. If possible we will collect

- patient had previous positive COVID-19 test (yes/no/unknown)
- type of test: PCR, point-of-care test
- date of test

Some patients consulting their GP may have had a SARS-CoV-2 infection in the past without having had a test. We will also ask:

- Have you ever been quarantined/asked to self-isolate due to being a contact of a confirmed COVID-19 case?
- If yes, date.

In the future we will also include results of antibody tests here.

Functional status

Low functional status is defined as needing help to bathe or to walk.

Health care utilization in the previous 12 months

In order to document and control for healthcare seeking behaviour in the control groups and the severity of underlying conditions, we will collect:

- the number of GP visits in the past 12 months before inclusion in the study
- the number of hospital admissions due to underlying conditions in the 12 months prior to inclusion in the study

Severity of illness episode

- As part of this illness episode, was this patient referred to hospital?
- Did the patient die as part of this illness episode?

Ethnicity

Ethnicity has been identified as a risk factor for COVID-19 in some studies. Collecting ethnicity data can be sensitive to collect and difficult to harmonise. This will be included as a pilot variable and revised in the future.

Exposure-related factors

The following questions (from contact to a confirmed case to protective factors) are entirely optional.

- Have you had close contact (<2 metres) with a confirmed COVID-19 case in the past 14 days?

  If yes, in which setting(s)?

  - home
  - education/school
  - workplace

- healthcare setting (GP, hospital)
- other (please specify)

- In the last 14 days, have you been out of the house and in environments of 2+ people not part of your household?

  If yes, which setting(s)? (select as many as appropriate)

  - home
  - education/school
  - workplace
  - public transport
  - healthcare setting (GP, hospital)
  - at a leisure setting (shopping, walk in parks, restaurants/bars/pubs, social gathering, sport, etc.)
  - other

- Household setting: how many members are in your household?

- Which precautions / protective measure have you taken in the past 14 days (select as many as are appropriate)?

  - face/nose/mouth protection (mask/scarf wearing)
  - extra handwashing
  - use of hand disinfectant
  - home disinfection
  - home isolation (I have not left my home)
  - social distancing
  - other

9. Data management

9.1. Data collection, entry and storage at site level

Web-based data collection methods or paper-based methods can be used. Double data entry is recommended unless electronic medical records are used.

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

Epiconcept provides the option of web-based data collection methods, if so desired by the sites: the Voozanoo web-based data entry platform, which is a secure system. These data can be accessed by the study site and the coordinating hub only. These methods can also be combined with paper-based methods. An example of a web or paper-based data collection form is in annex 7.

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.

Information on antigenic, when available, and genetic analyses can be stored separately on an Excel spreadsheet (see Annex 3).

All data should be stored and processed in a way compliant with GDPR.

- *Surveillance sites to specify procedures of data collection and entry.*
- *Surveillance sites to specify methods of data storage and their compliance with the GDPR requirements.*
- *Surveillance sites to provide a codebook that includes the variable names, variable descriptions, and the coding of variable values (see Annex 1).*
- *Surveillance sites to indicate all modifications in the variables collected during the course of the project.*

9.2. Data anonymisation and persistent unique identifier

All data sent from the sites should be anonymised. This means that the case-based data sent to the coordinating hub and the data on the Voozanoo data entry web platform should never include:

- Any names of patients
- Any addresses of patients
- Any medical registration numbers
- Any telephone numbers, email addresses or other contact details of patients
- Any dates of births (age in years is OK)
- Any other (combination of) information that increases the risk of identification

If these types of data are included in the data, the coordinating hub will not use them and delete them.

Each case-based record should have a unique identifier that the coordinating hub can use to identify a record when asking any questions to sites about data completeness or quality. This identifier should be persistent over the whole course of the surveillance/study (it should not change).

Each GP practice collecting data should have a unique and persistent identifier associated with it that the coordinating hub can use to identify a record when asking any questions to sites about data completeness or quality. This identifier should be persistent over the whole course of the surveillance (it should not change).

- *Surveillance sites to describe how and who performs the database anonymisation prior to local data analysis*

Aggregated data

Since aggregated data are counts per stratum, they contain no personal identifiers. However, if there are small numbers per cell in the denominator, then the combination of age group, sex and region may identify cases. As regional information is most useful at national level, sites can decide not to send information on the region to the coordinating hub or aggregate regions if small numbers by region are present.

- *Surveillance sites to specify procedures if small numbers are present in aggregated data.*

9.3. Data transfer, frequency of data transfer/reporting and storage at coordinating level

Enhanced surveillance (case-based) data

The frequency of reporting new data from study sites to the coordinating hub for surveillance data will initially be **monthly**. This may be revised to less frequent reporting according to COVID-19 incidence among sites participating and the recruitment strategy within primary care sites. This frequency will be reviewed on a regular basis.

Aggregated surveillance data

The frequency of reporting new aggregated data from study sites to the coordinating hub for surveillance data will initially be weekly. This may be revised to less frequent reporting according to COVID-19 incidence and the recruitment strategy within primary care sites. This frequency will be reviewed on a regular basis.

Potential data sharing with ECDC

In addition, ECDC is working with their legal department to identify the best way to share data with them. This protocol will be updated accordingly. For more information on this, data transfer, and storage at coordinating level, please see Annex 5.

9.4. Data checking and cleaning

Data checking will be carried out at site level, and also at pooled level by the coordinating team. Summary and frequency tables as well as visual representations of appropriate variables are used to find illegal, implausible or missing values within the dataset. Checks for inconsistencies are carried out (e.g. date of swabbing before date of onset of symptoms, or in the case of aggregated data, large differences between subsequent weeks in the denominators). These values should be checked against the questionnaires or queried with the GP. Any missing data will be described.

Any changes to the data during the cleaning process are documented and stored separately from the crude database. A guide and/or an example Stata do-file for data checking is provided if so desired.

At pooled level, questions arising after data checking will be queried with the sites using the unique identifiers (case-based data) or week numbers (aggregated data), so records can be traced back whilst maintaining anonymity. Data cleaning (recoding) will only take place in agreement with the site.

9.5. Site-specific reporting on enhanced surveillance data

Each individual study site will analyse and report their data. Some study sites will analyse the I-MOVE-COVID-19 surveillance data together with other surveillance data. Other study sites will analyse them separately.

Proportion of patients not consenting will be documented.

Surveillance participants will be described by baseline characteristics:

- sex
- age groups
- region

Surveillance participants will be described according to the following characteristics overall and by age group and sex (depending on sample size):

- time: new cases by week, month
- symptoms
- absence, presence of at least one, presence of more than one high-risk condition
- specific chronic conditions (e.g. respiratory, cardiovascular diseases)
- influenza and pneumococcal vaccination status
- respiratory co-infections
- referral to hospital or not

Each site can analyse the I-MOVE-COVID-19 data themselves. 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 regular pooled analysis will be carried out, for some surveillance indicators. The pooled analysis will be more relevant for the risk factor study and other studies. The regular pooled analysis will consist of a brief descriptive analysis.

A more detailed surveillance report will be produced at three specific time points, as outlined in section 10.5.

9.6. Site-specific reporting on aggregated surveillance data

Foremostly, the aggregated surveillance is important at national and subnational level, in order to measure the impact of/inform decisions on mitigation measures at a lower level geographical unit and to understand burden of disease. Each site should aim to report the results of their aggregated surveillance on a weekly basis through their own channels (e.g. sentinel bulletins, websites, etc.).

Each individual study site will document and report:

- number of participating GP practices and GPs
- the length of the surveillance period overall
- the proportion of patients consenting (source population)

Aggregated weekly numbers of suspected COVID-19 cases

Numbers of suspected COVID-19 patients, consulting in primary care, will be described for each week (depending on sample size), by:

- sex
- age groups
- by sex and age groups
- by region

If denominators are available, sites can calculate rates of suspected COVID-19 by region, age group and/or sex.

Aggregated weekly numbers of suspected COVID-19 cases, tests carried out and laboratory-confirmed COVID-19 cases

The analysis will be carried out as above and will also include for each week (depending on sample size):

- numbers of lab-confirmed COVID-19 cases (by age groups, sex, region)
- proportion of suspected COVID-19 cases tested (by age group, sex, region)
- proportion of laboratory-confirmed COVID-19 cases among tests (by age group, sex, region)

Note that the proportion of laboratory-confirmed COVID-19 cases among tests is approximate, as the laboratory results may not correspond exactly to the tests reported for a given week, due to a delay in receipt of laboratory results. However it is a good approximate indicator.

Each site can analyse the I-MOVE-COVID-19 data themselves. The coordinating hub can provide example scripts if desired or carry out the site-specific data analysis at the site's request.

9.7. Pooled analysis

The coordinating team will conduct the pooled analysis.

For **enhanced surveillance**, we will report:

- number of participating countries, GP practices and GPs
- the length of the surveillance period overall
- the number of patients captured by the surveillance: total, those consenting, by age-group and by participating site
- A descriptive analysis of COVID-19 cases by time, place and patient characteristics
- A descriptive analysis of SARS-CoV-2 clades by time, place and patient characteristics

The description of cases will include the distribution by age, sex, week/month of symptom onset, chronic conditions, co-morbid respiratory infections, pregnancy, smoking status, vaccination status (influenza and pneumococcal vaccines) and referral to hospital. This will be done among cumulative cases and among new cases. These variables will also be used in the study identifying risk factors for COVID-19.

For sites testing for other pathogens (multiplex RT-PCR), we will describe cases (co-infections) by pathogen identified.

More information on surveillance indicators can be found in annex 4.

For the pooled analyses of **aggregated surveillance** data, the coordinating team will report:

- number of participating countries, GP practices and GPs
- the length of the surveillance period overall
- the proportion of patients consenting
- a descriptive analysis of suspected COVID-19 cases by time, place and patient characteristics
- a descriptive analysis of laboratory-confirmed COVID-19 cases by time, place and patient characteristics

The description of cases will include the distribution by age group, sex, and possibly region. This will be done among cumulative cases and among new cases.

9.8. Dissemination of pooled results of enhanced surveillance

The coordinating hub aims at creating a short surveillance report at the frequency of data collection and transfer to the coordinating hub (initially defined at monthly, but to be reviewed regularly, see Annex 5).

The coordinating hub will only analyse and disseminate I-MOVE-COVID-19 results at site level (as opposed to a pooled analysis), if the site believes that this will provide an added value (e.g. to have all I-MOVE-COVID-19 country-specific data reported in the same place, or additional analyses carried out). Otherwise the coordinating hub will not disseminate the site level I-MOVE-COVID-19 data, but provide a link to a site-specific report or a website link. In some cases, I-MOVE-COVID-19 will do both (analyse the data and provide a site-specific report/website link). A brief descriptive analysis of pooled data will be part of the report.

The results will be placed on the I-MOVE-COVID-19 website (https://www.imoveflu.org/i-move-covid-19/) with unrestricted access. Some sites may prefer to embargo their data for a given period of time (e.g. 14 days), which will be noted in any surveillance analysis.

This report (in PDF) will also be uploaded onto the Zenodo platform as open access. Zenodo is a research repository launched in 2013 and hosted by CERN. It is GDPR-compliant and different access levels exist (https://about.zenodo.org/).

Additionally, the surveillance data underpinning the reports will be made publicly available on the Zenodo platform, along with a data codebook and scripts where possible. This will enable validation of the reports and ensure transparency and reproducibility. It will also enable other researchers to access and use the data for COVID-19 research. Site-specific data will only be shared openly with the site's consent.

More detailed surveillance reports will be produced at three specific time points, as outlined in section 10.5.

9.9. Dissemination of results of aggregated surveillance

The coordinating hub will only analyse and disseminate I-MOVE-COVID-19 results at site level (as opposed to a pooled analysis), if the site believes that this will provide an added value (e.g. to have all I-MOVE-COVID-19 country-specific data reported in the same place, or additional analyses carried out). Otherwise the coordinating hub will not disseminate the site level I-MOVE-COVID-19 data, but provide a link to a site-specific report or a website link. In some cases, I-MOVE-COVID-19 will do both (analyse the data and provide a site-specific report/website link). A descriptive analysis of pooled data will be part of the report.

The aggregated surveillance data and reports will be uploaded to the same platform for the less frequent (not weekly, see section 10.5) summary reports. This will be reviewed and any site-specific data will only be shared openly with the site's consent.

10. Further methods applicable to all surveillance

10.1. Ethical considerations

Each surveillance site will comply with national ethics committee requirements. Where required, informed consent will be sought from all participants or legal tutors. The national ethics committees will specify whether oral, written, or no consent will be required. A copy of the ethical approvals should be sent to the coordinating centre.

- ➢ *Each surveillance site to describe the procedures to comply with 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 surveillance site to send a copy of the ethical approval to the coordinating centre.*

10.2. Safety

During consultations and during the swabbing procedure, the safety of the practitioners is paramount. Any person swabbing, handling swabs and swabbing material, also in laboratories, should ensure that adequate personal protective equipment is used and hygiene measures followed.

- ➢ *Each surveillance site to state the safety measures carried out.*

10.3. 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). Surveillance site coordinators can decide which scientific conferences will be attended in order to present the results. An article presenting the results of the pooled analysis and 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 surveillance sites at the beginning of the study.

I-MOVE-COVID-19 results will be shared widely with other H2020 project teams and the public, as required by the European Commission's H2020 "open data" policy.

10.4. Training

Investigators and data collectors will be trained on the surveillance protocol before the start of the surveillance. They will receive the protocol and questionnaires.

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

10.5. Logistical aspects

Surveillance site leader

In each surveillance site, a principal investigator will coordinate the surveillance at the country level and act as focal point. The coordinating team is in charge of the pooled analysis.

The National Centre for Microbiology, Instituto de Salud Carlos III, Madrid, Spain is in charge of compiling and summarising the genetic data from the study sites.

Human resources

In each site, an investigator will be in charge of monitoring data collection at the GP office level. GPs will collect the information among consulting patients. The specific human resources needed in each country are detailed in the study annexes. Nivel ensures the overall coordination of the various surveillance sites.

Supervision

Site visits and joint workshops (remote if required) will be organised by the coordinating team/surveillance 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.

Reports

There are three surveillance bulletin reports that are deliverables for the I-MOVE-COVID-19 project. These should include data from all sentinel sites. They are due for submission at month 6, 12 and 24 of the project, corresponding to the 15th of September 2020, the 15th of March 2021 and the 15th of March 2022.

10.6. Limitations

The study includes only cases that are consulting GPs with COVID-19-like symptoms. Containment and mitigation strategies for the COVID-19 pandemic may differ by country depending on the case

management strategy (e.g. recommendation of contacting a specific COVID-19 helpline, or consulting a GP or health centre by telephone first). In some cases, the management strategy will have an impact on which cases consult a GP and are swabbed. This also may have an impact on the time lag between onset and respiratory specimen collection, and currently we do not know if this may affect positivity rates between surveillance sites. Beside the collection of the aforementioned data in the protocol, case-containment/ mitigation / health care seeking strategies should be described for each country.

For weekly aggregated data, the week of report of suspected COVID-19 cases will not correspond to the week of report of lab-confirmed COVID-19 cases, as there may be a delay for laboratory confirmation. Therefore the proportion of patients positive among tested will not be the exact proportion.

- *Each surveillance site to describe the potential limitations in terms of representativeness of the subjects included.*

11. References

[1] European Centre for Disease Prevention and Control (ECDC), World Health Organization Regional Office for Europe (WHO/Europe), 'Rapid risk assessment: Coronavirus disease 2019 (COVID-19) pandemic: increased transmission in the EU/EEA and the UK – eighth update', European Centre for Disease Prevention and Control. Accessed: Apr. 22, 2020. [Online]. Available: https://www.ecdc.europa.eu/en/publications-data/rapid-risk-assessment-coronavirus-disease-2019-covid-19-pandemic-eighth-update.

[2] M. Valenciano, B. Ciancio, and I-MOVE study team, 'I-MOVE: a European network to measure the effectiveness of influenza vaccines', *Euro Surveill.*, vol. 17, no. 39, Sep. 2012.

[3] European Centre for Disease Prevention and Control., 'Strategies for the surveillance of COVID-19', European Centre for Disease Prevention and Control. Accessed: Apr. 14, 2020. [Online]. Available: https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-surveillance-strategy-9-Apr-2020.pdf.

[4] World Health Organisation, 'World Health Organization. Laboratory biosafety guidance related to coronavirus disease 2019 (COVID-19): Interim guidance', 2020. Accessed: Apr. 24, 2020. [Online]. Available: https://apps.who.int/iris/bitstream/handle/10665/331138/WHO-WPE-GIH-2020.1-eng.pdf.

Annexes

Annex 1: List of variables, definitions and coding - minimum dataset I-MOVE-COVID-19 primary care-based enhanced surveillance

The following table contains a list of variables collected for I-MOVE-COVID-19 enhanced surveillance at primary care level. The variables for the minimum dataset are in blue. Additional variables for the surveillance are suggested in green. Variables only collected at the study site, but never transferred to the central data hub are in red.

The table in Annex 2 presents additional variables that can be collected for the risk factor study.

Sites can follow this variable naming and coding, or are welcome to code variables and values in their own way and send a codebook along with their data.

Variable name	Type	Values and coding	Definition
participate	Numeric (binary)	0 = No 1 = Yes	Agrees to participate
refuse	Text		Reasons for refusal to participate
id	Type of variable at discretion of site	[needs to be unique]	Unique and persistent identifier for each record
age	Numeric (continuous)	Integer	Age of each participant in years
sex	Numeric (binary)	0 = female 1 = male	Sex of study participant
[gppostcode]	Text		Postcode of GP practice
hcw	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Patient is a healthcare worker
height	Numeric (continuous)		Height in cm [6]
weight	Numeric (continuous)		Weight in kg [6]
bmi	Numeric (continuous)		Body Mass Index (kg/m2) [6]
obese	Numeric (categorical)	0 = No (BMI <30) 1 = BMI ≥30-39 2 = BMI ≥40 8 = Do not know	If BMI or height and weight are not collected: BMI ≥30-39; ≥40 [6]
onsetdate	Date	dd/mm/yyyy	Date of onset of symptoms
swabdate	Date	dd/mm/yyyy	Swabbing date
hosp_refer	Numeric (binary)	0 = No 1 = Yes	Referral of patient to hospital for COVID-19
fever	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Fever
malaise	Numeric (categorical)	0 = No	Malaise

[6] If possible, we suggest to collect either BMI as a continuous variable (bmi variable) OR height and weight. If none of these are possible then we suggest to use the categorical variable "obese" (page 33).

Variable name	Type	Values and coding	Definition
		1 = Yes 8 = Do not know	
myalgia	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Myalgia
cough	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Cough
sorethroat	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Sore throat
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
shortness of breath	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Weakness
anosmia	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Anosmia (Loss of sense of smell)
ageusia	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Aguesia/disguesia (Loss or distortion of sense of taste)
fatigue	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Fatigue
coryza	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Coryza
chestpain	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chest pain
chills	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chills/feverishness
nausea	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Nausea
vomiting	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Vomiting
diarrhoea	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Diarrhoea
stomachache	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Stomach ache
rash_derm	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Rash or other dermatological manifestation
conjunctivitis	Numeric (categorical)	0 = No	Conjunctivitis

Variable name	Type	Values and coding	Definition
		1 = Yes 8 = Do not know	
dizziness	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Dizziness
palpitations	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Palpitations
lab_sarscov2	Numeric (categorical)	0 = Negative 1 = Positive 2 = Inconclusive 8 = Do not know	Laboratory result for SARS-CoV-2 (positive/negative)
swab_type	Numeric (categorical)	1 = Nose 2 = Throat 3 = Both nose and throat 8 = Do not know	Type of swab taken
test_type	Numeric (categorical)	1 = PCR 2 = Point of care 3 = Other 8 = Do not know	Type of test used (if other, please specify)
lab_flu	Numeric (categorical)	0 = Negative 1 = Positive 2 = Not done 8 = Do not know	Laboratory result for influenza (positive/negative)
lab_rsv	Numeric (categorical)	0 = Negative 1 = Positive 2 = Not done 8 = Do not know	Laboratory result for RSV (positive/negative)
lab_metapneum	Numeric (categorical)	0 = Negative 1 = Positive 2 = Not done 8 = Do not know	Laboratory result for human metapneumovirus (positive/negative)
lab_rhinovirus	Numeric (categorical)	0 = Negative 1 = Positive 2 = Not done 8 = Do not know	Laboratory result for rhinovirus (positive/negative)
lab_adenovirus	Numeric (categorical)	0 = Negative 1 = Positive 2 = Not done 8 = Do not know	Laboratory result for adenovirus (positive/negative)
lab_bocavirus	Numeric (categorical)	0 = Negative 1 = Positive 2 = Not done 8 = Do not know	Laboratory result for bocavirus (positive/negative)
lab_seascorona	Numeric (categorical)	0 = Negative 1 = Positive 2 = Not done 8 = Do not know	Laboratory result for seasonal coronavirus (positive/negative)
genetic_group	Text		Laboratory result: SARS-CoV-2 genetic group
fluvaccany	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received flu vaccination in current season
pneumovacc	Numeric (categorical)	0 = No 1 = Yes	Received pneumococcal vaccination

Variable name	Type	Values and coding	Definition
		8 = Do not know	
diabetes	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Diabetes and endocrine disease
heart_dis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Heart disease
immuno	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Immunodeficiency and organ transplant
lung_dis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Lung disease
cancer	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Cancer
asthma	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Asthma
renal_dis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Renal disease
liver_dis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Liver disease
pregnant	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Pregnancy status
smoking		0 = Never 1 = Former 2 = Current 9 = Do not know	Never, former (stopped smoking at least 1 year before inclusion in the study), current smoker. Any smoking can be included: cigarettes, cigars, vaping, etc.
travel	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Travelled in the past 14 days?
travel_country	Text		Country of travel

Annex 2: List of variables, definitions and coding - Additional variables for I-MOVE-COVID-19 primary care-based study on risk factors for COVID-19

Variable name	Type	Values and coding	Definition
[gppostcode]	Text		GP postcode. This can be used for determining deprivation score and urban/rural. If not available, the "urban" variable below can be used. Sites should translate gppostcode to urban/rural and deprivation score at site level and do NEVER send GP postcode information to the central hub.
urban	Numeric (categorical)	0 = Predominantly rural 1 = Intermediate 2 = Predominantly urban 8 = Do not know	This variable can code for urban/rural or postcode of GP practice as a proxy
severity	Numeric (count)	integer	Number of hospitalisations previous 12 months for the chronic disease
gpvisit	Numeric (count)	integer	Number of GP consultations previous 12 months
fs_bath	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Requires assistance to bath
fs_walk	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Requires assistance to walk
antivir	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Administration of antivirals prior to swabbing
antivirdate	Date	dd/mm/yyyy	Date administration antiviral
antivirtype	Text		Type of antiviral (brand name)
res_home	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Living in a residential home
contra	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Contra-indication for influenza vaccination
fluvaccdate	Date	dd/mm/yyyy	Influenza vaccination date
fluvacctype	Text		Type of vaccine (brand name)
pneumovacctype	Numeric (categorical)	1 = PPSV23 2 = PCV13 3 = Other (pls specify) 8 = Do not know	Type of pneumococcal vaccine
pneumovacctype_other	Text		Other type of pneumococcal vaccine if not PPSV23 or PCV13
pneumoyear	Number		Year of receipt of pneumococcal vaccination
bcgvacc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Ever received BCG vaccine

Variable name	Type	Values and coding	Definition
bcgyear	Number		Year of receipt of BCG vaccination
statin	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Patient was under statin treatment before symptom onset
ace	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Patient takes angiotensin-converting enzyme inhibitors before symptom onset
arb	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Patient takes angiotensin II receptor blockers before symptom onset
nsaids	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Patient takes non-steroidal anti-inflammatory drugs before symptom onset
corticosteroids	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Patient takes corticosteroids
dmards	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Patient takes biological disease-modifying anti-rheumatic drugs
chemo	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Current/recent cancer chemotherapy
antithrom	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	antithrombotic/ platelet aggregation inhibitors
metformin	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Metformin before symptom onset
ethnicity	Numeric (categorical)		Collection to be determined according to country-specific guidelines
contact_case	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Optional: Patient has had close contact to a lab-confirmed COVID-19 cases in the following setting 14 days before their symptom onset
contact_case	Numeric (categorical)	1=home setting 2=education/school/daycare 3=workplace 4=healthcare setting 5=other	Optional: Setting in which patient has had close contact to a lab-confirmed COVID-19 cases in the following setting 14 days before their symptom onset
exposure	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Optional: In the last 14 days have you been out of the house and in environments of 2+ people?

Variable name	Type	Values and coding	Definition
exp_setting	Numeric (categorical)	1=home setting 2=education/school/daycare 3=workplace 4=healthcare setting 5=public transport 6=leisure setting (shopping, walk in parks, restaurants/bars/pubs, social gathering, sport, etc.) 7=other	Setting of recent exposure
household	Numeric (integer)	Integer	How big is household?
protection	Numeric (categorical)	1=face/nose/mouth protection (mask/scarf wearing) 2=extra handwashing 3=use of hand disinfectant 4=home disinfection 5=home isolation (I have not left my home) 6=social distancing 7=other	Which precautions have you taken in the past 14 days?
outcome	Numeric (categorical)	1=Hospitalised 2=Death 3=Hospitalisation + death 8=Unknown	As part of this illness episode, was the patient hospitalised or did the patient die?
prev_test	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Has the patient had a positive SARS-CoV-2 test prior to this illness episode?
date_prev_test	Date	dd/mm/yyyy	Date of previous SARS-CoV-2 test
type_prev_test	Numeric (categorical)	1 = PCR 2 = Point of care 3 = Other 8 = Do not know	Type of previous SARS-CoV-2 test
result_prev_test	Numeric (categorical)	0 = negative 1 = positive 8 = do not know	Result of previous SARS-CoV-2 test

Annex 3: Genetic and antigenic analysis data (examples)

The minimum amount of data needed to obtain genetic information from GISAID (sequences of all viruses should be sent to GISAID's open access EpiCoV platform) is country, I-MOVE-COVID-19 ID number and GISAID accession number. Additional information on CT value and selection for characterisation and reasons for not characterising can be additionally collected (see Table 2).

Table 2: Example of simplified data collection form for genetic data.

	Country	I-MOVE-COVID-19 ID number	GISAID accession ID number	Selected for characteri-sation?	Reasons for not characteri-sing?	CT value	Type of sample (primary specimen or isolate)
Strain 1							
Strain 2							

Where not all viruses were attempted to be sequenced, but only a random selection of them, additional information on sampling fraction should be provided, In order to better understand how viruses were selected for sequencing over time. An example can be seen in table 3.

Table 3: Example of documenting outlining how viruses were selected for sequencing over time

Time period	**First date of time period**	**Last date of time period**	**Sampling fraction used**	**Date used for definition of time unit (onset date, swab date, other)**	**Comments**
1					
2					
Example1	*01/10/2020*	*31/12/2020*	*1*	*Date of onset*	*(this is only an example; all specimens were characterised)*
Example2	*01/01/2021*	*15/02/2021*	*0.2*	*Date of onset*	*(this is only an example; 20% of all specimens were characterised)*

Annex 4: Surveillance indicators for enhanced surveillance (individual data)

1. Description of participating GP practices
    a. Number and % of participating GP practices
    b. Catchment population (if known) for participating GP practices, by country and pooled overall
    c. Map of participating GP practices
2. Description of cases (cumulative and by month), overall and by sex and age-group
    a. Number of laboratory-confirmed COVID-19 patients
3. Description of cases by clinical characteristics (overall and by age-group and sex)
    a. Number and % of laboratory-confirmed COVID-19 patients by
        i. symptoms
        ii. co-infections (if available)
        iii. GP referral to hospital
    b. Median length (and interquartile range) between onset of symptoms and swabbing
4. Description of laboratory-confirmed COVID-19 patients by preventive or risk factors
    a. Number and % of laboratory-confirmed COVID-19 patients (overall and by age-group and sex) by
        i. chronic conditions (presence/absence and by individual condition)
        ii. pregnancy status (among women)
        iii. smoking status (among those aged 15 and over)
        iv. influenza, pneumococcal vaccination
        v. recent travel
5. Laboratory indicators
    a. Number of cases by test performed? (if several tests are used?)
    b. Number of cases by clade (overall, by age-group, sex, country and region)

Annex 5: Data transfer, frequency of data transfer and data storage at pooled level

Software

For the multi-centre pooled analysis, surveillance sites will send an anonymised database to the coordinating team through the secure data transfer system EpiFiles (https://epifiles.voozanoo.net), which is a web platform which allows secure file exchanges between entities. Each site has a login and password for the EpiFiles system. Only the coordinating hub will be able to access the site-specific files.

Frequency

The frequency of reporting new data from study sites to the coordinating hub for surveillance data will initially be monthly for the individual level enhanced surveillance. This will be revised to less frequent reporting according to COVID-19 incidence and the recruitment strategy within primary care sites. This frequency will be reviewed on a regular basis. For the aggregated surveillance, the reporting of new data from study sites to the coordinating hub will be weekly.

For sites using the Voozanoo platform, data will be downloaded on a monthly basis for the individual level enhanced surveillance and weekly for the aggregated surveillance.

Study period of data to be transferred for individual level enhanced surveillance

Sites can send only new data to the coordinating hub each month, which will then be appended to previous data, or, if they prefer, they can send all data from surveillance start.

For some surveillance data there may be some changes to previous data (e.g. missing data completed, changes after data quality checks), therefore we recommend sending all data from surveillance start with each monthly transfer.

Data storage at pooled level

Please see also the I-MOVE-COVID-19 data management plan for more information (link to be provided soon).

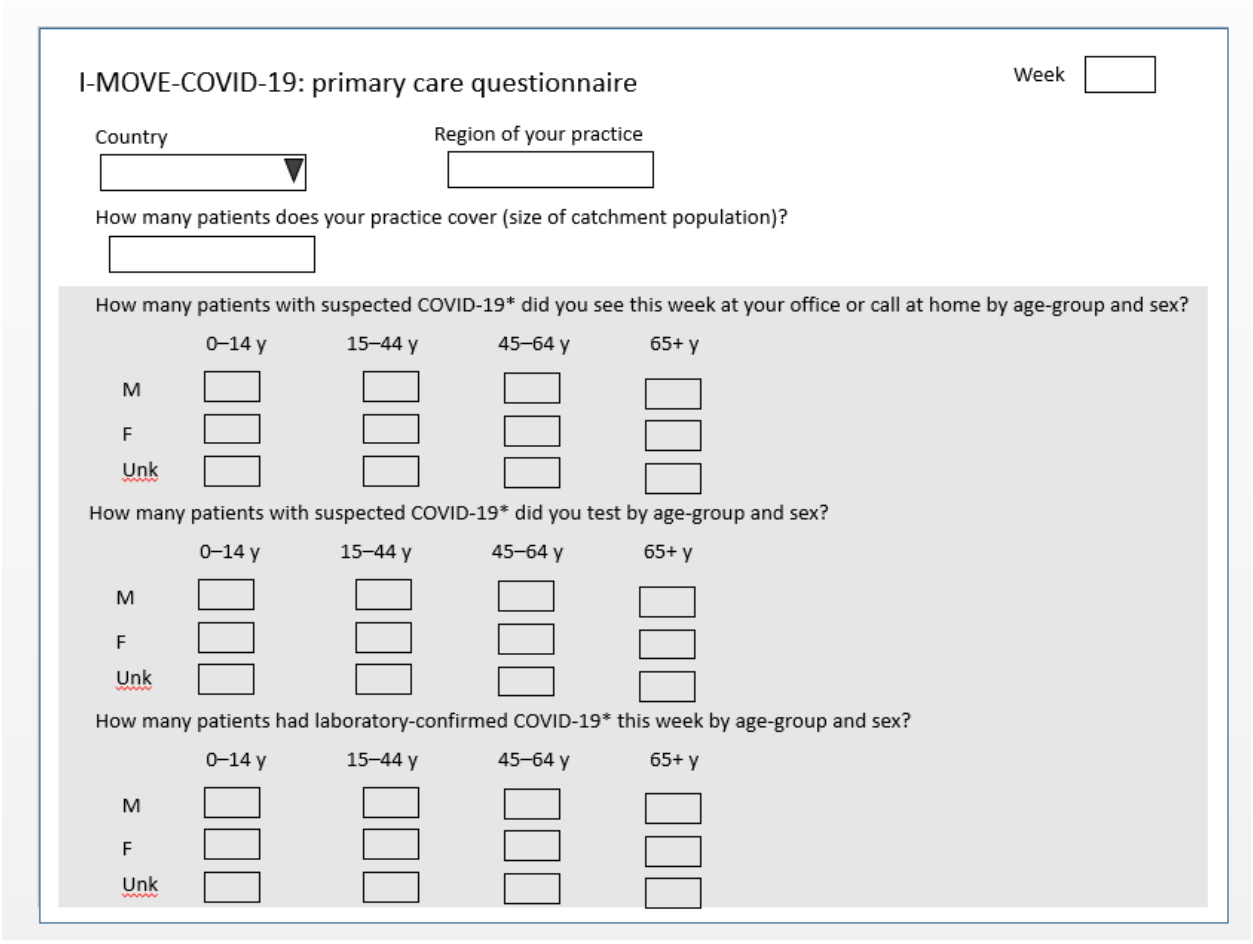
All anonymised data received from study sites will be stored in a GDPR-compliant manner. Work package leaders and the coordinators will have access to the pooled data. The pooled data will be stored in G Suite (provided by Google). This environment is GDPR-compliant and secure and private: https://gsuite.google.com/security/?secure-by-design_activeEl=data-centers

Annex 6: Study-specific annexes

Study specifications for each country are summarised in the annexes. Each surveillance site annex should include:

- description of the primary care practices participating in the study (number of GP practices, number of GPs, sampling strategy (all, systematic sample), information on sampling (face-to-face, self-swabbing, use of point-of-care tests, lack of PPE), catchment population)
- date of first case
- list of variables collected (and coding if different from suggested coding)
- pandemic (when applicable) vaccines used
- vaccine status ascertainment method
- details on methods for data collection, data entry and data transmission
- 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 GPs.

Annex 7: Mock up of aggregated surveillance data collection.



I-MOVE-COVID-19: primary care questionnaire

Week ☐

Country ▼ Region of your practice ☐

How many patients does your practice cover (size of catchment population)? ☐

How many patients with suspected COVID-19* did you see this week at your office or call at home by age-group and sex?

	0–14 y	15–44 y	45–64 y	65+ y
M	☐	☐	☐	☐
F	☐	☐	☐	☐
Unk	☐	☐	☐	☐

How many patients with suspected COVID-19* did you test by age-group and sex?

	0–14 y	15–44 y	45–64 y	65+ y
M	☐	☐	☐	☐
F	☐	☐	☐	☐
Unk	☐	☐	☐	☐

How many patients had laboratory-confirmed COVID-19* this week by age-group and sex?

	0–14 y	15–44 y	45–64 y	65+ y
M	☐	☐	☐	☐
F	☐	☐	☐	☐
Unk	☐	☐	☐	☐