# Study Protocol

## *Title: SARS-CoV-2 infection among health care professionals: demographic characteristics and serological and immune responses related to progression's phenotype (ProHEPiC-19)*

**PI and Study coordinators**: Violán-Fors, Concepción; Torán-Monserrat Pere;

Co_PIs: Garcia-Prado, Julia; Martínez-Cáceres, Eva

**Steering Committee**: Violán Fors, Concepción; Torán-Monserrat Pere; Garcia-Prado; Julia; Martínez-Cáceres, Eva

.

**Clinical trials protocol** id: 4r20-105

March, 18 2021

## INDEX

1. HEADER....4

2. ABSTRACT....5

3. STUDY GLOSSARY:....7

4. BACKGROUND....8

5. HYPOTHESES AND OBJECTIVES....10

5.1 HYPOTHESES....10

5.2 OBJECTIVES....10

5.2.1 PRIMARY AIM....10

5.2.2 SECONDARY AIM....11

6. METHODS....11

6.1. DESIGN:....11

6.2. POPULATION AND SETTING....11

6.2.1. SELECTION CRITERIA:....11

6.3. STUDY PERIOD....11

6.4. SAMPLE SIZE (POWER)....12

6.5. PROCEDURES....12

6.5.1. RECRUITMENT....13

6.6 VARIABLES MAIN COVARIABLES....14

6.6.2. OUTCOME MEASURES....16

6.7.DATA SOURCES....18

6.8 STATISTICAL ANALYSIS....18

6.8.1 STATISTIC TO MEET PRIMARY AIM: TO CONSOLIDATE A PROSPECTIVE COHORT OF HEALTH CARE WORKERS (HCW) TO GENERATE EPIDEMIOLOGICAL AND CLINICAL HIGH-QUALITY DATA.....18

6.8.2. STATISTICS TO MEET SECONDARY AIMS....18

6.9 ETHICS....20

7. REFERENCES....21

8.CONTRIBUTIONS OF AUTHORS AND AUTHOR AFFILIATIONS....22

9.MODIFICATION OF THE PROTOCOL....22

10.DISSEMINATION .....22

11. PROTOCOL AND REGISTRATION .....23

12.DATA STORAGE AND SECURITY .....23

13. ACKNOWLEDGEMENTS.....24

14. COMPETING INTERESTS.....24

15. LANGUAGE.....24

16. COMPENSATION .....24

17. OPEN ACCESS.....25

18.PATIENT CONSENT DETAIL .....25

19. PROVENANCE AND PEER REVIEW NOT COMMISSIONED; EXTERNALLY PEER REVIEWED.....25

20. DATA SHARING STATEMENT.....25

21. APENDICES.....25

## 1. HEADER

TITLE OF THE PROPOSAL: SARS-CoV-2 infection among Health Care Workers: demographic characteristics and serological and immune responses related to progression's phenotype (ProHEpiC-19)

**PIs and Coordinators**: Pere Torán Monserrat & Concepció Violán Fors

**Co-PIs**: Julia Garcia Prado; Eva Martínez Cáceres

**Steering Committee**: Violán Fors, Concepción; Torán Monserrat Pere; Garcia Prado, Julia; Martínez-Cáceres, Eva.

**ORIGINAL PROTOCOL** (Written in English in March, 31 2020)

Project members/ Contributions:

**Germans Trias i Pujol Research Institute (IGTP).** Concepció Violán Fors, Pere Torán Monserrat, Ester Lucas Varas, Aleix Pujol Gimeno

**Institut Universitari d'Investigació en Atenció Primària Jordi Gol (IDIAP Jordi Gol)**. Noemi Lamonja Vicente, Carla Chacón Valera, Francesc Ramos Roure, Aleix Pujol Gimeno, Josep M Manresa Dominguez, Albert Roso Llorach, Lucia Carrasco Ribelles, Marta Soldevilla Garcia, Ester Moral Roldan, Magda Alemany Costa, Eva Olivares Ortega, Monica Monteagudo Zaragoza, Pilar Montero Alia, Marta Bujalance Devesa, Mariella Soto Espinoza, Cristina Palacios Fernandez, Rosalia Dacosta Aguayo.

**Institut de recerca de la Sida. IrsiCaixa**: Julia Garcia Prado, Athina Kielpilanen , Oscar Blanch Lombarte, Miguel Angel Marin Lopez, Julieta Carabelli, Ruth Peña Poderós, Esther Jimenez Moyano, Dan Ouchi Vernet, Eulalia Grau Segura.

**Atenció primària Metropolitana Nord de Barcelona. Institut Català de la Salut**. Nuria Prat Gil, Josep Mº Bonet Simó, Nemesio Moreno Millán, Ester Badia Perich, Mónica Piña Rodriguez, Elena Domenech Graells, Eduard Moreno Gabriel, Victòria Sabaté Cintas, Mª Jose Argerich González, Assumpció Vazquez Duran, Alex Ortega Roca, Anna Devesa Pradells, Antonio Negrete.

**Hospital Universitari Germans Trias i Pujol. Institut Català de la Salut** Eva Martínez Cáceres, Bibiana Quirant Sánchez, Gema Fernandez Rivas, Sonia Molinos Abos, Jaume Barallat Martinez de Osaba, Bibiana Quirant Sanchez, Joan Matllo Aguilar, Elena Domenech Graells, Fernando Armestar Rodriguez, Maria Doladé Botias.

**Projected timing :** May, 30 2020- 30 Juny 2022

**Funding bodies:** Departament de Salut. Generalitat de Catalunya . Call COVID19-PoC number SLT16_04,l and Ministerio de Ciencia e Innovación . Convocatoria COVID-10 2020. Number COV20/00660

## 2. ABSTRACT

Introduction: Coronavirus Disease 2019 (COVID-19) has caused a global pandemic. Epidemiological and clinical inter-individual differences, symptomatology, recovery and humoral response against Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) are key factors to better understand and predict the course of the pandemic. As Health Care Workers (HCWs) are caring for infected patients they are more susceptible to infection, which not only is critical for their own health but also because it results in a shortage of HCWs that seriously affects health services. Thus, maintaining the health and welfare of HCWs and enabling their rapid return to work is vital to overcome this crisis. The ProHEpiC-19 cohort presents data on the immune response of HCWs infected with SARS-CoV-2. This dynamic cohort was started in March 2020 and still continues including participants.

Objectives:

Primary: To consolidate a prospective cohort of Health Care Workers (HCWs) to generate epidemiological and clinical high quality data. This information will be relevant to improve health policies and clinical COVID-19 protocols. This cohort will also be used as an ongoing platform to implement SARS-CoV-2 research projects with particular emphasis on incidence rate, reinfection, vaccines, and long term immune response.

Secondary:

1. To determine the kinetics of SARS-CoV-2 antibodies and cellular immune response in early, mid, and long periods of immunization.

2. To assess the relation between clinical variables and initial RT-PCR results with the interindividual differences in the immune response in early, mid, and long periods of immunization.

3. To analyze differentially expressed cytokines as biomarkers of disease progression in early, mid, and long periods of immunization.

Methods and analysis: Longitudinal, dynamic, prospective cohort study with a 12-month follow-up, which is being conducted in 4 primary-care centres and one hospital of Northern Metropolitana Nord of Barcelona (Spain). For now, the study consists of 1350 participants divided into 2 cohorts: 1) Healthy-Exposed HCWs: 675 not infected by SARS-CoV-2 (RT-PCR with a negative result and negative SARS-CoV-2 antibodies at baseline) and 2) Infected HCWs: 675 symptomatic participants (those with new persistent cough, temperature ≥37.5°C, anosmia, or ageusia or other compatible symptoms with COVID-19) or asymptomatic participants diagnosed by positive RT-PCR test and/or SARS-CoV-2 antibodies (IgM, IgG at baseline). Primary outcomes include: humoral and cellular immune response, quantitative antibodies to SARS-Cov-2, SARS-CoV-2 antibody levels related to progression phenotype, clinical spectrum of SARS-Cov-2, symptomatology, demographics and other variables that may be predictive of immune response.

Follow-up: baseline, 15 days, 1, 3, 6, 9 and 12 months.

Findings to date: Current literature has shown that the immune response is maintained for a minimum of 2 months. Nevertheless little is known about the association between the immune response and the progression phenotype of COVID-19 .

Future plans: This prospective cohort offers the possibility to study associations between immune response and progression phenotype according to age and gender as well as long-term immune response. In turn, we will be able to examine possible cumulative effects, taking into account several

clinical variables. The study is ongoing and we plan to extend it to increase the size of the cohort until 2024.

Ethics and dissemination: The study is approved by the Ethics Committee (IDIAPJGol ref 20/067-PCV IGTP ref COV20/00660 (PI-20-205). Dissemination will occur through presentations at National and International conferences and publications in international peer-reviewed journals.

Trial registration number: ID: 4R20-105

Keywords: SARS-CoV-2; COVID-19; Cohorts; Epidemiology, SARS-CoV-2 antibodies, Primary Health Care, Hospital, Health Care Workers, Immunology Response COVID-19.

## 3. STUDY GLOSSARY:

| | |
|---|---|
| Coronavirus Disease 2019 | COVID-19 |
| Severe acute respiratory syndrome coronavirus | SARS-CoV-2 |
| Health Care Workers | HCWs |

## 4. BACKGROUND

COVID-19, caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), was declared pandemic in March 2020 by the WHO [1]. At present (March, 30, 2020), the virus has infected more than 58.537 people worldwide with an associated case mortality rate of 1% to 15%, depending on the country and age [2].

Particularly, COVID-19 has affected a very high percentage of Health Care Workers (HCWs). In fact, 26% of all individuals who tested positive for SARS-CoV-2 by PCR in Spain, where transmission of SARS-CoV-2 has been intense, were HCWs [3]. So, HCWs are a high-risk group for developing the infection, that is exceptionally relevant for three reasons: 1) because of the clinical manifestations, severe in some cases, produced by the disease itself; 2) because as HCWs are caring for infected patients, they are more susceptible not only to infection, but also to infecting co-workers, patients and family members; and 3) because infection in this group leads to absenteeism from work, reducing the number of professionals available for health care. This situation results in a shortage of HCWs that seriously affects health services. Thus, maintaining the health and welfare of HCWs and enabling their rapid return to work is vital to overcome this crisis. Furthermore, as HCWs are constantly exposed to the virus, they are an interesting group to investigate the kinetics and antibody production in early, medium and long-term immunisations against SARS-CoV-2

In fact, a major current question is if individuals who have been infected with SARS-CoV-2, whether symptomatic or asymptomatic, knowingly or unknowingly, develop permanent immunity to the virus. Answering this question is crucial for epidemiologists, policy makers, health service professionals and society at large. To do so, three different types of diagnostic tools can be used: 1) RNA-based tests to determine whether viral genes are present in nasal and throat swabs, indicating active infection (by RT-PCR); 2) Detection of antibodies against SARS-CoV-2 in plasma or capillary blood. Antibodies can also indicate active infection but, more importantly, they indicate if the individual has already been infected and is able to develop a satisfactory immune response based on the presence of these antibodies. 3) Cytokine storm testing is a more exploratory diagnostic tool that can include assessment of the presence of proinflammatory cytokines in response to SARS-CoV-2 infection. The presence of a 'cytokine storm' has been linked to disease severity and progression and to a particular feature of the pathogenic features of SARS-CoV-2 infection [4].

Different studies have shown that levels of antibody response differ depending on the clinical spectrum of SARS-CoV-1 [5]. Nevertheless, little is known regarding immune response to SARS-CoV-2 It has been observed that the highest antibody responses are presented by people with a more severe clinical spectrum, hospital admissions and intensive care units. The highest response has been shown by IgM and IgG against SARS-CoV-2 antigens [6,7]. Despite this, there are still few studies that have been able to accurately analyze the results of the immune response over a period of 12 months. However, the immune response to SARS-CoV-2 infection is heterogeneous, with clinical spectra ranging from individuals with asymptomatic infections to individuals with very severe clinical manifestations that can even lead to death. It is possible that some individuals will never develop sterilizing immunity following infection with SARS-CoV-2 or that multiple exposures will be needed for affinity maturation and development of long-lasting protection. All these points have to be explored and they are extremely important for re-vaccination criteria.

At the current stage of the pandemic, laboratory tests measuring antibody response and seroconversion remain essential. Studies on indicators of inflammatory response and the prognostic value of other Covid-19 indicators and acute phase reactants are also needed. In addition, the identification of pro-inflammatory biomarkers may have added value in disease prognosis and provide novel information for

the therapeutics of immunosuppression of the inflammatory process in the acute phase. Furthermore, by conducting serological studies of the population we can achieve the following objectives: 1) To obtain a dynamic, qualitative and quantitative study of the immune response to SARS-CoV-2; 2) To accurately define the infection rate of a specific affected area, which is essential to determine the mortality rate of the infection; 3) To identify individuals with strong antibody responses, who could be donors to generate convalescent serum therapies; and 4) To determine who has and who lacks immunity against the disease. This is crucial especially because this serological information can be used to strategically deploy HCW who are immune to infection, thus limiting the risk of exposure and inadvertent spread of the virus.

As mentioned above, HCWs are a high-risk group for acquiring SARS-CoV-2 infection. Some of these professionals with symptoms or with high-risk contacts are confined and unable to return to their clinical activities while awaiting the results of the (RT)-PCR tests. It should be noted that all of them are confined to their homes for a minimum period of 8 days depending on the PCR result at 8, 14 and 28 days. This is causing a shortage of HCWs, which is seriously affecting the health services and putting additional pressure on a sector already overburdened by the pandemic.

Since the initial symptoms are very non-specific and compatible with any other respiratory virus, it is very difficult to screen for other diseases. It is therefore essential to identify factors associated with the protection of professionals against SARS-CoV-2. This is essential because the evolution of the SARS-CoV-2 pandemic makes it likely that new epidemic waves will appear within a short period of time. To a large extent, the factors conditioning the evolution of the infection, the permanence of the virus and the immunity of infected persons and that of healthy carriers of the infection have not yet been clearly identified.

To date, seroprevalence studies conducted in Spain have used cross-sectional designs. Thus, they cannot provide systematic data on long-term antibody and cellular immune responses.

Thus, thorough serological and prospective studies are needed to determine whether ADE manifests among SARS-CoV-2 infections, either because of prior homologous infection or cross-reactive antibodies from other HCoVs. This will have particular relevance for vaccination plans.

Another interesting point that deserves further analysis is analyzing the different antibody specificities generated and their dynamics. Several studies have examined various types of antibodies identified against SARS-Cov-2 (Spike1, Spike2, Nucleocapside, envelopes), identifying variable responses [8,9]. However, little is known about the simultaneous kinetics of the different antibodies over time. This can be analyzed with our cohort and also with the antigens used in the various vaccines that will be in the market in the next months. Thus, given the virus's short evolution, it is necessary to create an exhaustively monitored cohort. This cohort will make it possible to identify the kinetics of SARS-CoV-2-specific antibodies in individuals with different degrees of severity of COVID-19 and to know which levels of antibodies are protective against new possible reinfections in 3 years.

Crucially, this serological information will be used to strategically deploy HCW immune to the infection, limiting the risk of exposure and inadvertent propagation of the virus.

In conclusion, the present study is crucial to obtain epidemiological and clinical data to improve health policies for the prevention of SARS-CoV-2 infection. In addition, the study will be conducted simultaneously with the course of the pandemic, which will make it possible to obtain real-time data on the different waves that occur and, in turn, to study the SARS-CoV-2 endemic. Therefore, ProHEpIC-19 can be used as an ongoing platform to implement SARS-CoV-2 research projects with particular emphasis on incidence rate, reinfection, vaccines, and long term immune response.

# 5. HYPOTHESES AND OBJECTIVES

## 5.1 HYPOTHESES

The specific hypotheses are:

Infection with SARS-CoV-2 will produce different levels of immune response with respect to:

1. The kinetics and immune response is different depending on the clinical spectrum of COVID-19 disease.

2. The immune response is conditioned by age, sex and other clinical variables

3. Cytokines are good biomarkers to identify COVID-19 disease progression.

## 5.2 OBJECTIVES

### 5.2.1 PRIMARY AIM

To consolidate a prospective cohort of Health Care Workers to generate epidemiological and clinical high quality data. This information will be relevant to improve health policies and clinical COVID-19 protocols. This cohort will also be used as an ongoing platform to implement SARS-CoV-2 research projects with particular emphasis on incidence rate, reinfection, vaccines, and long term immune response.

Primary Outcome Measure:

1. Creation prospective cohort of health care workers

Include 675 exposed HCW participants and 675 infected HCW participants againts SARS-CoV-2, cohorts will be compared at each time point in terms of sociodemographic, epidemiological, clinical, and immunological information available. an exploratory bivariate analysis will be performed using the tests of Chi Square, ANOVA, Kruskall-Walis, depending on the application conditions assumptions.

Time Frame: Baseline, to 12 months after the beginning of the study

2. Cohort description demografics ( age, sex, academic level, housing characteristics, work variables )

Descriptive analysis of the participants will be performed using the number and percentage for categorical variables, and mean and standard deviation or median and quartiles 1 and 3 for quantitative variables, an exploratory bivariate analysis will be performed using the tests of Chi Square, ANOVA, Kruskall-Walis, depending on the application conditions assumptions.

Time Frame: Baseline, to 12 months after the beginning of the study

2. Cohort description clinical spectrum (asymptomatic, mild-moderate Illness, severe-critical)

Cohort comparison, an exploratory bivariate analysis will be performed using the tests of Chi Square, ANOVA, Kruskall-Walis, depending on the application conditions assumptions.

[Time Frame: Baseline, to 12 months after the beginning of the study]

### 5.2.2 SECONDARY AIM

1. To determine the kinetics of SARS-CoV-2 antibodies and cellular immune response in early, mid, and long periods of immunization.

2. To assess the relation between clinical variables and initial RT-PCR results with the interindividual differences in the immune response in early, mid, and long periods of immunization.

3. To analyze differentially expressed cytokines as biomarkers of disease progression in early, mid, and long periods of immunization.

## 6. METHODS

### 6.1. DESIGN:

Dynamic prospective study conducted in two cohorts of health care workers (HCW) (healthy and infected). Each cohort will include 675 participants.

Biological, clinical, psychological and social warning signs follow-up closely monitored during one year.

### 6.2. POPULATION AND SETTING

This is a multicenter study (public primary care centers set in Mataró, Sabadell and Santa Perpètua and the public third level hospital Hospital *Germans Tries i Pujol* located in Badalona) with HCW cohorts recruited from Gerència Territorial Metropolitana Nord of the Catalan Institute of Health (ICS), which consists of 7,776 HCW, including physicians, nurses, COVID-19 researchers, medical residents and other essential workers in direct contact with patients during present and future pandemic waves.

### 6.2.1. SELECTION CRITERIA:

HCWs (physicians, nurses, auxiliars, COVID-19 researchers and other essential workers in direct contact with patients during the first, second or third wave of COVID-19) at risk of infection due to direct patient contact. Random sampling is being used for the inclusion of subjects. Different recruitment strategies are being applied to contact with participants such as email, phone calls.

Inclusion criteria:

≥ 18 years of age

Accept to take part in the study and sign the informed consent

To be a HCWs infected or exposed to SARS-CoV-2.

Exclusion criteria:

Not accepting taking part in the study and/or not to sign the informed consent

Not to be a HCW exposed to SARS-CoV-2

### 6.3. STUDY PERIOD

From March, 31 2020 until recruitment is finished (estimated on June, 15 2021).

## 6.4. SAMPLE SIZE (POWER-)

According to estimates on March, 23 2020 1350 people meet eligibility criteria.

We will require 675 participants in each cohort to analyse the effects of SARS-CoV-2 on immunity. The sample will be representative of the study population.

*Sample Size Estimation*

In a one-way ANOVA study, sample sizes of 522, and 522 are obtained from the 2 groups whose means are to be compared. The total sample of 1044 subjects achieves 90% power to detect differences among the means versus the alternative of equal means using an F test with a 0.05 significance level. The size of the variation in the means is represented by the effect size $\eta^2 = \sigma m^2 / (\sigma m^2 + \sigma^2)$, which is 0.01 [10-13]

*Dropout-Inflated Sample Size*

| Average Group Sample Size n | Group | Dropout Rate | Sample Size Ni | Dropout-Inflated Enrollment Sample Size Ni' | Expected Number of Dropouts Di |
|---|---|---|---|---|---|
| 522 | 1-2 | 10% | 522 | 580 | 58 |
| | Total | | 1044 | 1160 | 116 |

| Average Group Sample Size n | Group | Dropout Rate | Sample Size Ni | Dropout-Inflated Enrollment Sample Size Ni' | Expected Number of Dropouts Di |
|---|---|---|---|---|---|
| 522,00 | 1 - 2 | 10% | 522 | 580 | 58 |
| | Total | | 1044 | 1160 | 116 |

## 6.5. PROCEDURES

Participants entering the study after meeting exclusion/exclusion criteria (Table 1), give written informed consent (informed consent and patient information sheet, see Appendix 1) l and they are assigned to a cohort (Figure 1) . Cohorts are divided into:

1) Healthy pre-cohort: HCW exposed to but not infected with SARS-CoV-2, negative PCR-test or Ig M (N) and IgG (N)

2) Non-recent Infected cohort: HCW infected with SARS-CoV-2 (>15 days RT-PCR, positive)

3) Recent Infected cohort infected HCW : HCW with acute SARS-CoV-2 infection (<15 days RT-PCR, positive)

All participants undergo a standardised follow-up according to the specific cohort to which they are assigned (Table 2). At study entry, we record the following information: baseline socio-demographic

data, clinical data (symptoms), epidemiological data (contact with a positive and symptoms of COVID-19) (table 3), information on the work environment, data on the diagnosis of COVID-19 (previous biological samples taken and dates of sick leave due to COVID-19), use of personal protective equipment, previous health status and habits. At subsequent visits, epidemiological data, clinical data (symptoms) and vaccinations are recorded.

Finally, initial participant's cohort assignation is reviewed and reclassificated on basis on laboratory results (RT-PCR test or SARS-CoV-2 antibodies IgM and IgG against the nucleocapsid (N)) in two final cohorts:

1) Health Care Workers Healthy Cohort (HCW_H): This cohort includes HCW with negative laboratory test RT-PCR and SARS-CoV-2 antibodies IgM(N), and IgG(N).

2) Health Care Workers Infected Cohort (HCW_I): This cohort includes HCW with positive laboratory test (RT-PCR and/or SARS-CoV-2 antibodies IgM and IgG) This cohort is subdivided into three categories according to the illness severity:

Asymptomatic or Presymptomatic Infection: Individuals with positive laboratory evidence of COVID-19 but no COVID-19 signs nor symptoms.

Mild-Moderate Illness: Individuals with positive laboratory evidence of COVID-19 who have any of the various signs and symptoms of COVID-19 (e.g., fever, cough, sore throat, malaise, headache, muscle pain, nausea, vomiting, diarrhea, loss of taste and smell) but who do not have shortness of breath, dyspnea, or abnormal chest imaging and home health care.

Severe-Critical Illness: Individuals with positive laboratory evidence of COVID-19 who have any of the various signs and symptoms of COVID-19 (e.g., fever, cough, sore throat, malaise, headache, muscle pain, nausea, vomiting, diarrhea, loss of taste and smell) and that have shortness of breath, dyspnea or abnormal chest imaging, treated in the hospital with or without admission to the ICU.

This whole process is being conducted by trained multidisciplinary research staff (family physicians, nurses, assistant manager and project manager). See Figure 1

### 6.5.1. RECRUITMENT

The Basic Prevention Unit of Northern Metropolitan Area of Barcelona (Primary Health Care and Hospital), Spain (NM_BCN) offers HCW to participate in ProHEPIC-19 study, if they accept; the researchers of ProHEpiC-19 contact these participants.

Eligible participants interested in the study, are accurately informed about the ProHEpiC-19 protocol, and signed informed consent to participate.

Participants complete several clinical questionnaires and immune clinical determination at baseline, 7 days, 15 days, 3, 6, 9 and 12, 16, 20, 24 months after the beginning of the study. (see Table 2)

Participant data files will be stored in numerical order (code system) and in a secure and accessible place and manner. Participant files will be maintained in storage for a period of 15 years after completion of the study.

Anonymised demographic data will be collected for those participants who decline to take part in the project, to assess the external validity of the recruited sample of participants.

## 6.6 VARIABLES MAIN COVARIABLES

1) Biological sex measured as a dichotomous variable in two categories (Male, Female)

2) Age (years): registered and analyzed as a continuous variable

3) Academic level: measured as a 5 levels variable (lower primary, completed primary, secondary, Post-secondary non-tertiary education, Short-cycle tertiary education, university).

4) Housing characteristics (meters, number of bedrooms, number of people in the same house, relationship, age of them)

5) Work variables (professional category, schedule, workplace's characteristics)

6) Clinical Spectrum of COVID-19 defined on procedure for action against cases of infection with the new SARS-CoV-2 coronavirus, defined according to national and international guidelines ( asymptomatic, Mild-Moderate Illness, severe-critical ) as defined in section 6.5

7) Clinical monitoring of COVID-19 symptoms, days of duration and monitoring Long Covid symptoms. (Table 3 , Table 4)

8) The date of symptom onset shall be considered as the date of appearance of the first symptoms. These symptoms, which are described in Table 2, are the confirmation of a positive RT-PCR test or antigen test (IgM (N) or IgG (N)).

9) Diagnostics. In the electronic medical records (EMR) for Primary Care and Hospitals, the diseases are codified using the International Classification of Diseases, version 10 (ICD-10), including the specific codes for COVID-19

10) Use of personal protective equipment (working and in the free time)

11) Medication prescribed. Classification number of drugs according to the Anatomical Therapeutic Chemical (ATC) Classification System.

12) Hospital admissions: admissions and emergency procedures according to the codes ICD-10 used in CMBD-HA.

13) Comorbidities: diabetes, hypertension, ischaemic heart diseases, other chronic comorbidities

14) Mortality by all causes. The initial data for the calculation of risk of death will be determined from March 15, 2020.

15) Psychosocial and healthy habits changes during pandemic (dream, eating, sexuality, physical activity and familiar relationship)

16) Laboratory Assays

A)RT-PCR will be used as diagnostic test every 2 weeks RNA for RT-PCR tests will be extracted from fresh samples using the STARMag 2019-nCoV kit (Seegene, CA, USA) on a liquid-dispensing robot (STARlet Hamilton, USA). To inactivate the virus and extract viral RNA, 500 µl of the transfer medium containing the swab is combined with 500 µl of lysis buffer of the STARMag kit in a BSL-2 laboratory; following 10' of incubation at room temperature, 350 µl was used for nucleic acid extraction into a volume of 100 µl of extracted genome, from which 8 µl was taken for the 2019-nCoV PCR assay.The detection of RNA will be performed using the Allplex™ SARS-CoV-2 Assay is a multiplex real-time PCR

assay to detect 4 target genes of SARS-CoV-2 in a single tube. The assay is designed to detect RdRP, S and N genes specific for SARS-CoV-2, and E gene for all of Sarbecovirus including SARS-CoV-2 (Seegene, CA, USA) according to the manufacturer's instructions.

B) Quantification of antibodies to SARS-CoV-2. We performed a previous validations study with 6 different IVD-CE approved ELISA tests and selected the commercially available test kits - the AntiSARS-CoV-2 IgG and IgM ELISA (Inmunodiagnostic). This ELISA test had a Sensitivity 92.5% and a specificity 93.3% for IGM and IgG against Nucleocapsid (N) and Spike (S) antigens. For this, the best reagents and the most relevant antigen specificities will be analysed. We will perform test at baseline, 7 days, 15 days, 3, 6, 9 and 12,16,20,24 months after the beginning of the study.

We conducted a pre-validation study (March-April 2020) with 6 different IVD-CE approved ELISA tests and selected commercially available Anti-SARS-CoV-2 IgG and IgM ELISA kits (Inmunodiagnostic Limited ©). These ELISA tests had a sensitivity of 92.5% and a specificity of 93.3% for IGM and IgG against the nucleocapsid (N) antigen. Positivity thresholds were provided by the assay manufacturers and were considered positive with an index value greater than 1.1, indeterminate from 0.9 to 1.1 and negative if <0.9 index units.

Positive participants with Rt-PCR + or IgM (N) or IgG (N) were also tested for antibodies against the Spike (S) subunit of SARS-CoV-2. This was done by enzyme immunoassay for the qualitative determination of IgG class antibodies against COVID-19 (SARS-CoV-2) in human serum with DECOV1901. This reagent complies with the technical specifications for sensitivity, specificity, and standardisation according to WHO International Standards and Reference Panel for anti-SARS-CoV-2 antibody (WHO/BS.2020.2403). Positivity thresholds were provided by the assay manufacturers and were considered positive with an index value greater than 40, indeterminate from 32 to 40 and negative if <32 Ul/ml

Positive participants for serology or/and RT-PCR were also tested for antibodies against the Spike (S) subunit of SARS-CoV-2. This was done by enzyme immunoassay for the quantitative determination of IgG class antibodies using DECOV1901 (Demeditec Diagnostics GmbH©). This reagent complies with the technical specifications for sensitivity, specificity, and standardisation according to WHO International Standards and Reference Panel for anti-SARS-CoV-2 antibody (WHO/BS.2020.2403). Positivity thresholds were provided by the assay manufacturers and were considered positive with an index value greater than 40, indeterminate from 32 to 40 and negative if <32 Ul/ml

C) Additional Immune response laboratory assays:

C1) Additionally, blood samples of HCW (infected and healthy participants) will be collected and stored for more exhaustive determinations of the immune response in later studies.

C2) INFg Elispot assays. To monitor both SARS-CoV-2 specific CD4+ and CD8+ T-cell responses we performed an IFNγ ELISPOT assay. The design compromises a comprehensive screening using peptide pools for the nucleocapsid, membrane, S1 and S2 regions of the spike protein described in Olvera et al., (2020) vaccines and total recombinant proteins for the nucleocapside and the spike. In brief, cells will be thawed and incubated overnight at 37°C in the presence of SARS-CoV2 peptide pools described (5ug/ml for each peptide) and PHA (15 µg/ml), as positive control. Finally, plates were revealed with biotinylated anti-human IFNγ, streptavidin-alkaline phosphatase and its coloured substrate (Mabtech, Sweeden).

IFNγ secreting cells were quantified under Immuno Capture and Immuno Spot software to calculate the number of Spot Forming Cells (SFC). Cryopreserved PBMCs will be stimulated overnight with CoV-2 peptide pools (5ug/ml for each peptide) and PHA (15 µg/ml), as the positive control. Spot Forming Cells (SFC) will be counted using an INFg-ELISpot kit (Mabtech), AP-conjugate kit (BioRad) and an automated ELISPOT Reader Unit (CTL) as previously described (Dalmau et al AIDS 2014). Wells will be considered positive if they contained at least 50 spot-forming cells per 106 PBMCs above the background level (2X mean + 3Xstandard deviation).

C3) Citokines multiplex assay. Cryopreserved plasma samples will be used in a 45-plex assay of soluble mediators. The plates will be read with a Luminex instrument (Luminex 200, Austin Luminex, USA) and data will be analysed using MILLIPLEX Analyst 5.1 software (Merck Millipore Darmstadt, Germany). Measures will be performed in duplicates.

C4) Combinatorial flow-cytometry of lineage, activation inducible markers and intracellular cytokines Cryopreserved PBMCs will be stimulated with reactive peptide pools identified by ELISpot and incubated overnight. In the ICS, we will combine T-cell lineage markers (CD3, CD4, CD8, CD45, CCR7, CD27), activation inducible markers (AIMs: OX40, CD25 and CD137) and cytokines (IL2, IFNg, TNF). Samples will be acquired on an LSR Fortessa cytometer (BD), data analysed using FlowJo software v10 and polyfunctional cytokine profiling represented using SPICE v6 (NIAID, NIH).

### 6.6.2. OUTCOME MEASURES

Outcome measures of immune response (SARS-CoV-2 Antibodies (N), cytokines, T-cells) and Covid-19 clinical symptoms will be measured at baseline and prospectively aiming to observe possible associations between possible prognostic factors and immune response. The baseline and prospective measurements will allow to assess their prognostic role and whether some of them appear as confounding factors.

**Primary Outcome Measure:**

1. Creation prospective cohort of health care workers

Include 675 exposed HCW participants and 675 infected HCW participants againts SARS-CoV-2, cohorts will be compared at each time point in terms of sociodemographic, epidemiological, clinical, and immunological information available. an exploratory bivariate analysis will be performed using the tests of Chi Square, ANOVA, Kruskall-Walis, depending on the application conditions assumptions.

[Time Frame: Baseline, to 12 months after the beginning of the study]

2.Cohort description demografics ( age, sex, academic level, housing characteristics, work variables )

Descriptive analysis of the participants will be performed using the number and percentage for categorical variables, and mean and standard deviation or median and quartiles 1 and 3 for quantitative variables, an exploratory bivariate analysis will be performed using the tests of Chi Square, ANOVA, Kruskall-Walis, depending on the application conditions assumptions.

[Time Frame: Baseline, to 12 months after the beginning of the study]

2. Cohort description clinical spectrum (asymptomatic, mild-moderate Illness, severe-critical)

Cohort comparison, an exploratory bivariate analysis will be performed using the tests of Chi Square, ANOVA, Kruskall-Walis, depending on the application conditions assumptions.

[Time Frame: Baseline, to 12 months after the beginning of the study]

**Secondary Outcome Measures:**

4. Kinetics of SARS-CoV-2. IgM Nucleocapside

IgM (nucleocapside) ELISA kits (Inmunodiagnostic Limited ©). Positivity thresholds were provided by the assay manufacturers and were considered positive with an index value greater than 1.1, indeterminate from 0.9 to 1.1 and negative if <0.9 index units

[Time Frame: Baseline, 7 days, 15 days, 3, 6, 9 and 12 months after the beginning of the study]

5. Kinetics of SARS-CoV-2. IgG Nucleocapside

IgG (nucleocapside) ELISA kits (Inmunodiagnostic Limited ©). Positivity thresholds were provided by the assay manufacturers and were considered positive with an index value greater than 1.1, indeterminate from 0.9 to 1.1 and negative if <0.9 index units

[Time Frame: Baseline, 7 days, 15 days, 3, 6, 9 and 12 months after the beginning of the study]

6. Kinetics of SARS-CoV-2. IgG Spike

IgG (spike). ELISA kits DECOV1901 (Demeditec Diagnostics GmbH©). Positivity thresholds were provided by the assay manufacturers and were considered positive with an index value greater than 40, indeterminate from 32 to 40 and negative if <32 UI/ml

[Time Frame: Baseline, 7 days, 15 days, 3, 6, 9 and 12 months after the beginning of the study]

7. Kinetics of SARS-CoV-2. T-Cell

SARS-CoV-2 specific CD4+ and CD8+ T-cell responses we performed an IFNγ ELISPOT assay. Wells will be considered positive if they contained at least 50 spot-forming cells per 106 PBMCs above the background level (2X mean + 3Xstandard deviation).

[Time Frame: Baseline, 7 days, 15 days, 3, 6, 9 and 12 months after the beginning of the study]

8. To assess the relation between clinical variables and initial RT-PCR results in the whole sample and by sex.

To study the differences between clinical spectrums and initial RT-PCR we will use ANOVAs or Kruskal-Wallis tests, after checking normality assumption using a Shapiro-test

[Time Frame: Baseline, to 12 months after the beginning of the study]

9. To analyse the relation between clinical variables and the interindividual differences in the immune response in early, mid, and long periods of immunization in the whole sample and by sex

To study the differences between clinical spectrums and immune response in early period we will use ANOVAs or Kruskal-Wallis tests, after checking normality assumption using a Shapiro-test . Similarly, to look for differences in antibody levels between sex, either a t-test or a Mann-Whitney test will be performed.

[Time Frame: Baseline, 7 days, 15 days, 3, 6, 9 and 12 months after the beginning of the study]

10. Cytokines as biomarkers of disease progression in early, mid, and long periods of immunization.

Cryopreserved plasma samples will be used in a 45-plex assay of soluble mediators. The plates will be read with a Luminex instrument (Luminex 200, Austin Luminex, USA).Appropriate statistical tests (i.e. t-test or Mann-Whitney to compare between sexes and ANOVA or Kruskal-Wallis to compare between clinical spectrums) will be used after checking for normality (Shapiro-test)

[Time Frame: Baseline, 7 days, 15 days, 3, 6, 9 and 12 months after the beginning of the study]

## 6.7.DATA SOURCES

We will simultaneously use various information sources

A) Database created for the register of ProHEpiC-19.

The data will be centralized in an internal centralized data warehouse from the Catalan Health Institute. After data collection and curation, the data is ready to be analyzed to (1) validate the accuracy on demographic, clinical, social, and psychological data of professionals; (2) register SARS-COV-2 antibodies; (3) results in RT-PCR test.

B) Primary care electronic health records (e-Cap) contain demographic data and clinical variables: health conditions (ICD-10 codes) and diagnostic date, medication prescribed, comorbidities, lab test results (including RT-PCR).

C) Case Report Form (Teleform) specific for the study, which will include a specific identifier for the patient and the serological and immunological tests' information. This information will be shared among the various sources of information about the study employing individual, anonymized participants' identification. Immunological data will be collected to understand the vaccine's protection and the relationship between immune response and reinfection.

## 6.8 STATISTICAL ANALYSIS

Data collection: The data will be exported to the statistical packages used in the study. The people that meet eligibility criteria will be identified and included in the study database. A comprehensive cleaning of the data will be conducted to eliminate inconsistencies, duplicities and treatment indications inconsistent with the diagnosis and analysis of missing data will also be conducted. Additionally, a longitudinal validation of diagnostics and medicines will be carried out.

### 6.8.1 STATISTIC TO MEET PRIMARY AIM: TO CONSOLIDATE A PROSPECTIVE COHORT OF HEALTH CARE WORKERS (HCW) TO GENERATE EPIDEMIOLOGICAL AND CLINICAL HIGH-QUALITY DATA.

A descriptive analysis of the participants will be performed using the number and percentage for categorical variables, and mean and standard deviation or median and quartiles 1 and 3 for quantitative variables. Frequency and percentage of followed-up patients and dropouts will be updated at each timepoint. Profiles of dropouts and visited people will be assessed using the appropriate statistical test in order to prevent potential biases due to drop out. Cohorts will be compared at each time point in terms of sociodemographic, epidemiological, clinical, and immunological information available. To analyze this, an exploratory bivariate analysis will be performed using the tests of Chi Square, ANOVA, Kruskall-Walis, depending on the application conditions assumptions.

### 6.8.2. STATISTICS TO MEET SECONDARY AIMS

#### 6.8.2.1. TO DETERMINE THE KINETICS OF SARS-COV-2 ANTIBODIES AND CELLULAR IMMUNE RESPONSE IN EARLY, MID, AND LONG PERIODS OF IMMUNIZATION.

Two methods will be followed to study the kinetics of SARS-CoV-2 antibodies since symptoms onset. Locally Stimated Scatterplot Smoothing (LOESS) models will be fitted as a first descriptive approach. Then, nonlinear mixed-effects (NLME) models will be fitted to get the curves that would determine the kinetics.

LOCALLY ESTIMATED SCATTERPLOT SMOOTHING (LOESS) MODELS

LOESS models will fit as an initial descriptive manner of studying the antibodies, both IgM and IgG, evolution since diagnosis day. LOESS regression consists of fitting simple models (e.g., first or second degree polynomials) in subsets of data determined by a nearest neighbours algorithm. These LOESS curves will best stratified either per clinical condition or gender, and will be shown together with their 95% confidence interval (CI). We will use all the data available except for those from participants who never tested positive, grouping the data points per participant in the figures.

NONLINEAR MIXED-EFFECTS (NLME) MODELS

Mixed-effects models are generally useful when there are multiple measures per participant, as in this study. Random effects are included in the model to test if they may be causing any significant, and generally unexpected, significant influence in the modelling built by the fixed effects. Therefore, introducing random effects in the model help to control unobserved variance.

We will use NLME models to study possible differences in the dynamics (kinetics) of SARS-CoV-2 antibodies (IgM, IgG) levels since symptoms onset. After fitting a general model, the differences will be studied between clinical spectrums and biological sex, setting the largest group as reference groups for modelling. To fit these models, we will use all the data available except for those from participants who never tested positive. Time after diagnosis was discretized as follows: Tests made in the first 14 days since diagnosis are treated as "Day 15". Tests from days 15 to 29, as "Day 30", days 30 to 59 as "Day 60", days 60 to 89 as "Day 90", days 90 to 179 as "Day 180", days 180 to 269 as "Day 270", days 270 to 360 as "Day 360", and days 360 up to end of follow-up as “Day 450”. However, antibody values beyond 450 days after diagnosis were not considered, as they were only available for 14 participants.

The general equations to be fitted will be:

$$IgM(t) = b_1 + (b_0 - b_1)e^{-k_1 t} - b_1 k_1 (e^{-k_1 t} - e^{-k_2 t})$$

$$IgG(t) = b_1 + (b_0 - b_1)e^{-k_1 t} - b_1 k_1 (e^{-k_1 t} - e^{-k_2 t})$$

Having $t$ normalized per the maximum number of days (360). This general equation corresponds to an equation describing an exponential rise followed by an exponential decay. $b_0$ corresponds to the baseline value, $b_1$corresponds to the asymptotic value (i.e., antibody levels at $t \to \infty$), $k_1$ corresponds to the rise rate, and $k_2$ to the decay rate.

Nonlinear mixed-effects models will be built using the nlme package [14].

#### 6.8.2.2. TO ASSESS THE RELATION BETWEEN CLINICAL VARIABLES AND INITIAL RT-PCR RESULTS WITH THE INTERINDIVIDUAL DIFFERENCES IN THE IMMUNE RESPONSE IN EARLY, MID, AND LONG PERIODS OF IMMUNIZATION.

The incidence by SARS-CoV-2 will be linked to the sociodemographic, epidemiological, clinical and immunological information collected in the project. To analyze this, an exploratory bivariate analysis will be performed using the appropriate bivariate tests. Chi-squared tests will be used to find if there were any significant differences between the expected and the observed frequencies in categorical comparisons (e.g., the prevalence of symptoms per sex or clinical spectrum).. To study the differences between clinical spectrums and initial RT-PCR we will use ANOVAs or Kruskal-Wallis tests, after checking normality assumption using a Shapiro-test . Post-hoc analyses will be performed to find significant differences between clinical spectrum and immune response. Similarly, to look for differences in antibody levels between sex, either a t-test or a Mann-Whitney test will be performed.

#### *6.8.2.3.* TO ANALYZE DIFFERENTIALLY EXPRESSED CYTOKINES AS BIOMARKERS OF DISEASE PROGRESSION IN EARLY, MID, AND LONG PERIODS OF IMMUNIZATION

Cytokine levels will be studied and compared between different clinical spectrum or sexes to determine their role as a biomarker of disease type. Appropriate statistical tests (i.e. t-test or Mann-Whitney to compare between sexes and ANOVA or Kruskal-Wallis to compare between clinical spectrums) will be used after checking for normality (Shapiro-test).Significance level will be set to 0.05. All the analyses will be performed using R version 3.6.3 or higher.

## 6.9 ETHICS

The biological samples for this investigation are available and stored under different IRB-approved study protocols and properly consented for investigational use. Study participants are informed about the objectives of the study and the interventions linked to their participation in it. They are also given a written information sheet detailing all the information about the project and asked to sign the informed consent form. The confidentiality and anonymity of the data are ensured in accordance with current state laws (Organic Law 3/2018, of 5 December, on Personal Data Protection and guarantee of digital rights). The study is being conducted in accordance with the articles pertaining to the Declaration of Helsinki, agreed at the 64th General Assembly in 2013. This project is under the tutelage of the Clinical Research and Ethics Committee (CEIC) of the Instituto Universitario de Investigación en Atención Primaria (IDIAP Jordi Gol) and incorporates its recommendations and suggestions.

Your samples may only be used for the purpose stated in this information sheet. Your samples may be used in ongoing studies of COVID-19 infection, but may also be retained for future research related to COVID-19 infection.

All studies using your samples must be approved by a Research Ethics Committee. You may request to be informed about the studies in which your samples have been used.

Your samples will be stored mainly in the Germans Trias i Pujol Health Sciences Research Institute (IGTP) Biobank, an institution that has been accredited by the health authorities *(Insitituto de Salud Carlos III. Ministerio de Ciencia e Innnovación*) specifically to store samples for research under the conditions required by law (number B.0000643) and IrsiCaixa biological collection “Pathogenesis of infectious diseases” with the registration number C.0006008. The treatment and use of all samples will be carried out in accordance with the Biomedical Research Act (14/2007) and Royal Decree 1716/2011, which regulates biobanks. The samples will be kept for 15 years. They may be transferred to third parties if the current Biobank legislation is complied with and they will not be sold under any circumstances.

In some cases, they may become part of sample collections accredited by the health authorities. In this case, they can neither be transferred to third parties nor sold, but they could later be integrated into the Biobank of our centre.

The samples will be identified in a coded form. This means that the information on the labelling of the samples cannot be linked to the person who donated the sample. Please see below the section on Confidentiality and data protection.

The biological samples for this investigation are available and stored under IGTP and IrsiCaixa specific research collection-approved study protocols and properly consented for investigational use. The study will be approved by the Ethics Committee (IDIAPJGol ref 20/067-PCV IGTP ref COV20/00660 (PI-20-205).

The study will be approved by the Ethics Committee (IDIAPJGol ref 20/067-PCV IGTP ref COV20/00660 (PI-20-205)

## 7. REFERENCES

1. https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov) ]

2. https://coronavirus.jhu.edu/us-map

3. Chen ZM, Fu JF, Shu Q, Chen YH, Hua CZ, Li FB, et al. Diagnosis and treatment recommendations for pediatric respiratory infection caused by the 2019 novel coronavirus. World J Pediatr.2020, http://dx.doi.org/10.1007/s12519-020-00345-5.

4. García-Sierra RM, Badia Perich E, Manresa Dominguez JM, Moreno Millan N, Sabaté Cintas V, Romero Martínez M, Moreno Gabriel E, Pera G, Seda Gombau G, Montellà Jordana N, Violan Fors C, Argerich González MJ, Bonet Simó JM, Prat Gil N, Torán Monserrat P. Descriptive study of the health service workers of a Primary Care Department confined by Covid-19.. Rev Esp Salud Publica. 2020 Sep 3;94:e202009106. Spanish. PMID: 32880381.

5. Liu W, Fontanet A, Zhang PH, et al. Two-year prospective study of the humoral immune response of patients with severe acute respiratory syndrome. J Infect Dis 2006;193:792-5.

6. Callow KA, Parry HF, Sergeant M, Tyrrell DA. The time course of the immune response to experimental coronavirus infection of man. Epidemiol Infect 1990;105:435-46

7. Choe PG, Perera RAPM, Park WB, et al. MERS-CoV Antibody Responses 1 Year after Symptom Onset, South Korea, 2015. Emerg Infect Dis 2017;23:1079-84.

8. Huang, A.T., Garcia-Carreras, B., Hitchings, M.D.T. et al. A systematic review of antibody mediated immunity to coronaviruses: kinetics, correlates of protection, and association with severity. Nat Commun 11, 4704 (2020).

9. Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, Adriano A, Beese S, Dretzke J, Ferrante di Ruffano L, Harris IM, Price MJ, Dittrich S, Emperador D, Hoof L, Leeflang MMG, Van den Bruel A. Antibody tests for identification of current and past infection with SARS-CoV-2. Cochrane Database of Systematic Reviews 2020, Issue 6.

10. Cohen, Jacob. 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates. Hillsdale, New Jersey

11. Desu, M. M. and Raghavarao, D. 1990. Sample Size Methodology. Academic Press. New York.

12. Fleiss, Joseph L. 1986. The Design and Analysis of Clinical Experiments. John Wiley & Sons. New York.

13. Kirk, Roger E. 1982. Experimental Design: Procedures for the Behavioral Sciences. Brooks/Cole. Pacific Grove, California.

14. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2020). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-144. Available at: https://CRAN.R-project.org/package=nlme

## 8.CONTRIBUTIONS OF AUTHORS AND AUTHOR AFFILIATIONS

Steering Committee: Violán Fors, Concepción; Torán Monserrat Pere; Garcia Prado, Julia; Martínez-Cáceres, Eva.

PIs and Coordinators: Pere Torán Monserrat & Concepció Violán Fors

Co-PI: Julia Garcia Prado; Eva Martínez Cáceres

Project members/ Contributions:

## 9.MODIFICATION OF THE PROTOCOL

Changes from the original plan will be summarized on a separate paper an create a new version of study protocol

## 10.DISSEMINATION

Core to the achievement of the increased awareness and impact of the SARS-CoV-2 immune response is the improvement of communication at every level and in particular between health professionals and patients.

In support of the project's objectives, a stakeholder communications strategy will be developed, which links the various groups to be engaged during the project with the most effective means of providing them with overall messages and ongoing engagement. The strategy will include:

a) Project visual identity

b) Key messages

c) Timing and content of research briefings and policy notes

d) Schedule of public activities.

At the same time, the project's and social media channels will be established. The strategy will detail specific outputs relating to the general media. The activities detailed in the strategy will be reviewed every six months.

Target groups for Dissemination and Communication

The communication activities to the target groups are outlined in the list below. Activities will target the following main stakeholders: (1) HCWs; (2) Health agencies, Health politicians; (3) Patients and patient organization.

a) Social networks: Over 100 stakeholders engaged at the month 12 ProHEpiC-19 project , and 1000 by the end of the project.

b) Briefing notes: At least six short, non-expert, briefing notes produced and distributed

c) Press Releases: At least three, non-expert, briefing notes produced and distributed

d) Publications and conference presentations: At least three scientific papers submitted and six conference presentations by the end of the project including a final conference.

Publications and conference presentations will be targeted according to the direction of the findings and to take advantage of complementary conference and journal special issue topics.

## 11. PROTOCOL AND REGISTRATION

This study is registered with ClinicalTrials.gov ID: 4R20-105

## 12.DATA STORAGE AND SECURITY

*1) Type and format of data collected/generated within the Project*

All data will be processed in a pseudo-anonymised way by coding the identifying variables in both the initial paper data collection booklet (CB) and the subsequently created database. Only the necessary data to achieve the research objectives will be collected. The necessary variables to conduct the project are: 1) 1) Biological sex; 2)Age (years); 3) Academic level; 4) Housing characteristics ; 5)Occupational situation ; 6) Clinical Spectrum of COVID-19; 7) Clinical monitoring of COVID-19 Symptoms; 8) Date first symptoms; 9) Diagnostics. In the electronic medical records; 10) Use of personal protective equipment; 11) Medication prescribed; 12) Hospital admissions; 13) Comorbidities; 14) Mortality by all causes; 15) Psychosocial and healthy habits changes during pandemic; 16) Laboratory Assays: RT-.PCR test, antibodies to SARS-CoV-2 ( IgM and Ig G against nucleocapside, IgG Against Spike) , INFg Elispot assays; Citokines multiplex assay; Combinatorial flow cytometry of lineage, inducible activation markers and intracellular cytokines. These variables can be consulted in the corresponding section of the methodology.

In accordance with Law 14/2007, the samples of each participant in the project will be collected directly from the person concerned and the procedure set out in the regulations on the use of biological samples will be followed to authorise their use. The samples will only be used for the project and will subsequently be destroyed.

*2) Procedure provided to access data (who, how and when can access), data ownership, repository to deposit data.*

-The necessary variables to perform the study will be obtained directly from the participants in the project by means of a self-completed form to be filled in by the participant. The exploratory variables, biological samples and complementary explorations will be obtained by accredited personnel. All variables and data obtained will require the prior consent of the participant, in accordance with the provisions of articles 6.1.a) and 9.2.a) of the RGPD, as well as additional provision 17.2.d of the LOPD-GDD.

The Catalan Institute of Health (ICS), the IDIAP Jordi Gol Research Institute and the Germans Trias i Pujol Research Institute (IGTP) act as data controllers in the framework of this observational study. The project database will be housed in the computers of the Unitat de Suport a la Recerca Metropolitana Nord of the ICS, which will act as data processor. Initially, international data transfer is not foreseen. If this were to take place, a new informed consent would be requested from the study participants.

*3) Procedure planned to ensure specific ethical and legal requirements.*

The database has been created and stored on a local computer where the organisation has installed the software and therefore only accessible on computers that have a trusted connection via VPN and secure credentials (certificates, RSA keys or complex passwords). This computer follows the security standards set by the ICS in compliance with current requirements.

## 13. ACKNOWLEDGEMENTS

The authors are grateful to the participants in this study. We also want to thank the Medical College of Barcelona, The Nurse College of Barcelona, Closca, Leroy Merlin, Barfim, Roche, for their personal protection equipment and health material donations.

In addition, the researchers would like to thank Núria Prat Gil, Director of Primary Care at the Metropolitana Nord, and Josep Ma Bonet Simó, Assistant Director of Primary Care at the Metropolitana Nord.This project was funded by Departament de Salut. Generalitat de Catalunya . Call COVID19-PoC number SLT16_04

## 14. COMPETING INTERESTS

The authors declare no competing interests.

## 15. LANGUAGE

Case report forms must be completed in Catalan and Spanish and Inform Consent and Sheet participant information is available in English ,also.

All written information and other material to be used by subjects and investigative staff must use vocabulary and language that are clearly understood.

## 16. COMPENSATION

Participants will be covered by usually public insurance. They do not receive any compensation

## 17. OPEN ACCESS

A public dissemination of the results will be made.

## 18.PATIENT CONSENT DETAIL

The informed consent and the patient information sheet are shown in appendix

## 19. PROVENANCE AND PEER REVIEW NOT COMMISSIONED; EXTERNALLY PEER REVIEWED.

Scientific Committee of The Foundation University Institute for Primary Health Care Research Jordi Gol i Gurina (IDIAPJGol)

## 20. DATA SHARING STATEMENT

The study data will not be openly available. If researchers or institutions are interested in the data, the authors will be able to release the data after asking the study participants for their consent and signing the relevant legal agreements allowing the data to be released to third parties.

## 21. APENDICES

The tables and figures cited in the text are shown below.

**Table 1. ProHEpiC-19 Selection criteria.**

| | |
|---|---|
| Inclusion criteria | ≥ 18 years of age<br><br>Accept to take part in the study and sign the informed consent according to the Declaration of Helsinki.<br><br>To be a health care professional worker infected or exposed to SARS-CoV-2. |
| Exclusion criteria | < 18 years old<br><br>Not to accept to take part in the study and/or not to sign the informed consent according to the Declaration of Helsinki.<br><br>Not to be a health care professional worker exposed to SARS-CoV-2 |

**Table 2 Participants assigned to each cohort**

| Visit | Pre-cohorts enrolled | Protocol |
|---|---|---|
| **0** | **All possible participants** | Telephone screening/personal interview to explain the study and to decide if is an eligible participant and pre-cohort |
| **1** | **Healthy exposed healthcare workers** | RT-PCR (SARS-CoV2)<br>Anti-SARS-CoV-2 IgG and IgM antibodies (Nucleopcapside)<br>Hemogram and serum biochemistry |
| | **Recent infected healthcare workers** | RT-PCR (SARS-CoV2)<br>Anti-SARS-CoV-2 IgG and IgM antibodies<br>Hemogram and serum biochemistry<br>T-cells and cytokines |
| | **No-recent infected healthcare workers** | No visit |
| **2** | **Recent infected healthcare workers** | Telephone symptoms control call |
| | **Healthy exposed healthcare workers and No-recent infected healthcare workers** | No visit |
| **3** | **Healthy exposed healthcare workers and No-recent infected healthcare workers** | No visit |
| | **Recent infected healthcare workers with positive RT-PCR in Visit 1 or previously** | With positive RT-PCR in Visit 1 or previously: RT-PCR (SARS-CoV2) and Anti-SARS-CoV-2 IgG and IgM antibodies |
| | | With negative RT-PCR in visit 1 and previously: Anti-SARS-CoV-2 IgG and IgM antibodies |
| **4** | **Healthy exposed healthcare workers** | Anti-SARS-CoV-2 IgG and IgM antibodies |
| | **Recent infected healthcare workers** | Anti-SARS-CoV-2 IgG and IgM antibodies<br>T-cells and cytokines |
| | **No-recent infected healthcare workers** | RT-PCR (SARS-CoV2) |

| | | |
|---|---|---|
| | | Anti-SARS-CoV-2 IgG and IgM antibodies<br>Hemogram and serum biochemistry |
| **5** | **Healthy exposed healthcare workers** | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| | **Recent infected healthcare workers with positive RT-PCR in Visit 1 or previously** | Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | **Recent infected healthcare workers with negative RT-PCR in visit 1 and previously** | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| | **Non-recent infected healthcare workers** | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| **6** | **Healthy exposed healthcare workers** | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| | **Recent infected healthcare workers with positive RT-PCR in Visit 1 or previously** | Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | **Recent infected healthcare workers with negative RT-PCR in visit 1 and previously** | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| | **Non-recent infected healthcare workers** | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| **7** | **Healthy exposed healthcare workers** | Anti-SARS-CoV-2 IgG and IgM antibodies |
| | **Recent infected healthcare workers with positive RT-PCR in Visit 1 or previously** | Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | **Recent infected healthcare workers with negative RT-PCR in visit 1 or previously** | Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | **Non-recent infected healthcare workers** | Anti-SARS-CoV-2 IgG and IgM antibodies |
| **8** | **Healthy exposed healthcare workers** | Anti-SARS-CoV-2 IgG and IgM antibodies |
| | **Recent infected healthcare workers with** | Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |

| | | |
|---|---|---|
| | **positive RT-PCR in Visit 1 or previously** | |
| | **Recent infected healthcare with negative RT-PCR in visit 1 and previously** | Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | **Non-recent infected healthcare workers** | Anti-SARS-CoV-2 IgG and IgM antibodies |
| **9** | **Healthy exposed healthcare workers** | Anti-SARS-CoV-2 IgG and IgM antibodies |
| | **Recent infected healthcare workers with positive RT-PCR in Visit 1 or previously** | Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | **Recent infected healthcare with negative RT-PCR in visit 1 and previously** | Anti-SARS-CoV-2 IgG and IgM antibodies, T-cells and cytokines |
| | **Non-recent infected healthcare workers** | Anti-SARS-CoV-2 IgG and IgM antibodies |

**Table 3 . Initial symptoms of COVID-19**

| **SARS-COV2 CLINICAL SYMPTOMATOLOGY** |
|---|
| Headache |
| Diarrhea |
| Dyspnea |
| Fever |
| Cough |
| Anosmia |
| Arthralgias |
| Asthenia |
| Shivers |
| Thoracic |
| Epigastralgia |
| Discomfort |
| Myalgias |
| Nausea |
| Odynophagie |
| Congestion |

**Table 4 Symptoms clinical monitoring***

| Symptom | Date of onset of the symptom | Duration (days) | Observations |
|---|---|---|---|
| Cardiovascular | | | |
| Fainting | | | |
| Pain/burning in chest | | | |
| Tachycardia | | | |
| Bradycardia | | | |
| Palpitations | | | |
| Visibly inflamed/ bulging veins | | | |
| Dermatologic | | | |
| COVID toe | | | |
| Dermatographia | | | |
| Other skin and allergy | | | |
| Peeling skin | | | |
| Petechiae | | | |
| Skin rashes | | | |
| Gastrointestinal | | | |
| Diarrhea | | | |
| Loss of appetite | | | |
| Vomiting | | | |
| Abdominal pain | | | |
| Nausea | | | |
| Constipation | | | |
| Gastroesophageal reflux | | | |
| Head/eyes/nose/ears/throat | | | |
| Runny nose | | | |
| Sore throat | | | |
| Hearing loss | | | |
| Other hearing/ear issues | | | |
| Tinnitus | | | |
| Vision symptoms | | | |
| Immunologic/autoimmune | | | |
| New allergies | | | |
| New anaphylaxis reaction | | | |
| Musculoskeletal | | | |
| Bone ache or burning | | | |
| Muscle aches | | | |

| | | | |
|---|---|---|---|
| Tightness of chest | | | |
| Joint pain | | | |
| Muscle spasms | | | |
| Neuropsychiatric | | | |
| Acute confusion/disorientation | | | |
| Changes to sense of smell and taste | | | |
| Dizzines, unsteadiness or balance issues | | | |
| Hallucinations | | | |
| Headaches and related symptoms | | | |
| Insomnia | | | |
| Other sleeping symptoms | | | |
| Sleep apnea | | | |
| Slurring eord/speech | | | |
| All neurological sensations | | | |
| Brain fog | | | |
| Memory issues | | | |
| Neuralgia (nerve pain) | | | |
| Speech/language issues | | | |
| Tremors | | | |
| Vibrating sensations | | | |
| Pulmonary/Respiratory | | | |
| Dry cough | | | |
| Rattling of breath | | | |
| Breathing difficulty (normal O2 saturation level) | | | |
| Cough with mucus production | | | |
| Coughing blood | | | |
| Other respiratory and sinus | | | |
| Shortness of breath | | | |
| Sneezing | | | |
| Reproductive/genitourinary/endocrine | | | |
| All menstrual issues | | | |
| Bladder control issues | | | |

| | | | |
|---|---|---|---|
| Systemic | | | |
| Elevated temperature (98.8-100.4F) | | | |
| Fever (>100.4F) | | | |
| Chills/flushing/sweats | | | |
| Fatigue | | | |
| Low temperature | | | |
| Other temperature issues | | | |
| Post exertional malaise | | | |

**Characterizing Long COVID in an International Cohort: 7 Months of Symptoms and Their Impact Hannah E. Davis1*, Gina S. Assaf1*, Lisa McCorkell1*, Hannah Wei1*, Ryan J. Low1,2*,Yochai Re'em1,3*,Signe Redfield1, Jared P. Austin4, Athena Akrami1,2*+ medRxiv preprint doi: https://doi.org/10.1101/2020.12.24.20248802*

**TABLE 5. Protocol to be followed according to the cohort of inclusion in the study**

| Visit | Study Cohorts | Protocol |
|---|---|---|
| 0 | All possible participants | Telephone screening/personal interview to explain the study and to decide if is an eligible participant and pre-cohort |
| 1 | Healthy Health Care Workers (HCW) | RT-PCR (SARS-CoV2)<br>Anti-SARS-CoV-2 IgG and IgM antibodies<br>Hemogram and serum biochemistry |
| | Infected HCW | RT-PCR (SARS-CoV2)<br>Anti-SARS-CoV-2 IgG and IgM antibodies<br>Hemogram and serum biochemistry<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | Healthy HCW | No visit |
| 2 | Infected HCW | Telephone symptoms control call |
| 3 | Healthy HCW | No visit |
| | Infected HCW with positive RT-PCR in Visit 1 or previously | With positive RT-PCR in Visit 1 or previously: RT-PCR (SARS-CoV2) and Anti-SARS-CoV-2 IgG and IgM antibodies |

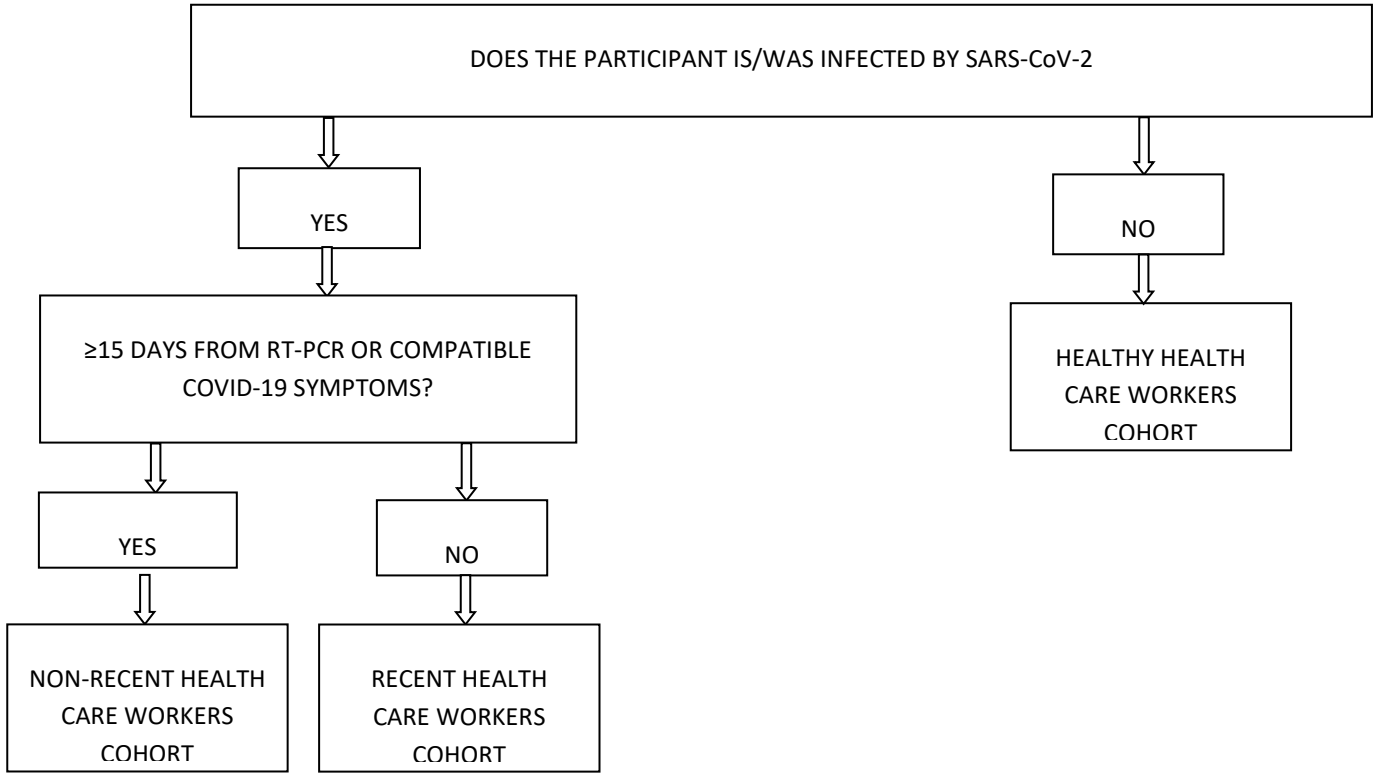
| | | |
|---|---|---|
| | | with negative RT-PCR in visit 1 and previously: Anti-SARS-CoV-2 IgG and IgM antibodies |
| | Healthy HCW | No visit |
| 4 | Healthy HCW | Anti-SARS-CoV-2 IgG and IgM antibodies |
| | Infected HCW | Anti-SARS-CoV-2 IgG and IgM antibodies<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | Healthy HCW | RT-PCR (SARS-CoV2)<br>Anti-SARS-CoV-2 IgG and IgM antibodies<br>Hemogram and serum biochemistry |
| 5 | Healthy HCW | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| | Infected HCW with positive RT-PCR in Visit 1 or previously | Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |

| | | |
|---|---|---|
| | Infected HCW with negative RT-PCR in visit 1 and previously | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| | Healthy HCW | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| 6 | Healthy HCW | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| | Infected HCW with positive RT-PCR in Visit 1 or previously | Anti-SARS-CoV-2 IgG and IgM antibodies<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | Infected HCW with negative RT-PCR in visit 1 and previously | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |

| | | |
|---|---|---|
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies |
| | Healthy HCW | with antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | | without antibodies in Visit 4: Anti-SARS-CoV-2 IgG and IgM antibodies<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| 7 | Healthy HCW | Anti-SARS-CoV-2 IgG and IgM antibodies |
| | Infected HCW with positive RT-PCR in Visit 1 or previously | Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | Infected HCW with negative RT-PCR in visit 1 and previously | Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | Healthy HCW | Anti-SARS-CoV-2 IgG and IgM antibodies |
| 8 | Healthy HCW | Anti-SARS-CoV-2 IgG and IgM antibodies |
| | Infected HCW with positive RT-PCR in Visit 1 or previously | Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |

| | | |
|---|---|---|
| | Infected HCW with negative RT-PCR in visit 1 and previously | Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | Healthy HCW | Anti-SARS-CoV-2 IgG and IgM antibodies |
| 9 | Healthy HCW | Anti-SARS-CoV-2 IgG and IgM antibodies |
| | Infected HCW with positive RT-PCR in Visit 1 or previously | Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | Infected HCW with negative RT-PCR in visit 1 and previously | Anti-SARS-CoV-2 IgG and IgM antibodies,<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |
| | Healthy HCW | Anti-SARS-CoV-2 IgG and IgM antibodies<br>T-cells (INFg Elispot assays, Combinatorial flow-cytometry)<br>Citokines multiplex assay |

**Figure 1.Decision criteria for assigning participants to cohorts**



# INFORMED CONSENT

Study title: **SARS-CoV-2 infection among healthcare professionals: demographic characteristics and serological and immune responses related to progression's phenotype.**

I, .....

I have read the foregoing information sheet to the patient and / or participant.

I have had the opportunity to ask questions about the study.

I have received enough information about the study.

I have talked to:

I understand that my participation in this study is completely voluntary.

I understand that I can withdraw from the study:

1. When I want it,
2. Without giving explanations, and
3. Without affecting the health care that I will receive.

In accordance with the provisions of Regulation (EU) 2016/679 of the European Parliament and the Council of April 27 on Data Protection (RGPD) and Organic Law 3/2018, of December 5, on data protection and guarantee of digital rights, I declare to have been informed of my rights, of the purpose of collecting my data and of the recipients of the information.

I consent voluntarily to participate as a participant in this research. Date (day/month/year) ……..

Name of Participant (Name and surname)

……

Signature of Participant

……

Name of the clinical researcher

……

Signature of the clinical researcher

……

## PATIENT/PARTICIPANT INFORMATION SHEET

### Introduction

This Information Sheet is for people who are invited to participate in the research entitled: **"SARS-CoV-2 infection among healthcare professionals: demographic characteristics and serological and immune responses related to progression's phenotype.".**

**Name of principal investigator:** Pere Torán Monserrat / Concepción Violan Fors
**Name of the organisation:** IGTP / IDIAPJGol

You will receive information about this research and will be invited to be part of it. Before you decide, you are welcome to talk to anyone you feel comfortable with about the research. If there is anything you don't understand, don't hesitate to ask and we will review the information.

### Research objectives

At the moment, healthcare professionals are a group at high risk of contracting SARS-CoV-2 infection. The evolution of the SARS-CoV-2 pandemic suggests that epidemic waves may recur within a short period of time since the factors that condition the evolution of the infection, the permanence of the virus and the immunity of infected individuals are largely unknown. In addition, healthy carriers are not clearly identified. The main objective of this research is to consolidate a cohort of health professionals exposed to SARS-CoV-2, to measure the serological and immunological characteristics that allow distinguishing symptomatic from asymptomatic carriers and to compare it with a cohort of patients treated in the same health professionals' workplaces.

### Type of research and procedures

This is a prospective study with three cohorts, two of them consisting of healthcare workers and one consisting of patients. This research will involve the collection of biological samples for subsequent analysis, as well as the collection of socio-demographic and quality of life data:

1. At least two swab samples shall be taken for the RT_PCR diagnostic test of oropharyngeal and nasopharyngeal exudates to detect the presence of the virus. The first sample shall be taken when included in the cohort, the second sample 14 days after inclusion in the cohort and it may be necessary to obtain some more samples until the absence of virus is verified.

2. Two capillary blood samples obtained by finger prick will be taken on the same days as the exudate test for RT_PCR. This test is performed to measure C-reactive protein in blood as a marker of the degree of inflammation present in the organism.
3. Six blood samples will be taken for serological and immunological tests obtained by venous puncture. The first one at inclusion in the cohort, the second on day 14 after the first RT-PCR, the third at day 60, the fourth 90 days after the first RT-PCR test, the fifth between days 180 and 270

and the sixth at the end of the study, i.e. one year after the inclusion visit.

4. A questionnaire on demographic, clinical, social and psychological data will be administered during the first month after inclusion in the study. Additionally, some of the participants will be contacted afterwards with the sole purpose of offering them the opportunity to complement the information on the psychosocial impact of SARS-CoV-2 by means of an individual interview. This would only be carried out after obtaining a new ad-hoc informed consent, designed and validated by the reference Research Ethics Committee (REC).

**Selection of participants**

We are inviting 600 healthcare workers included in the protocol for assistance to people in isolation due to SARS-CoV-2 exposure, plus 300 additional health professionals with different risks of infection, systematically selected from the Northern Metropolitan Territorial Management of the Catalan Health Institute (ICS).

**Voluntary participation**

Your participation in this research is completely voluntary. It is entirely your choice whether or not to participate.

**Benefits**

The results of this study will contribute to the characterisation of an algorithm to identify carrier status and immune response in healthcare professionals with various levels of exposure to SARS-CoV-2. This will optimise the management of human resources with sound criteria for the safety of professionals and patients, thus strengthening the health system to cope with this pandemic and the new waves predicted. You will have the results of the tests and will be able to know your degree of immunisation against the virus.

**Processing of biological samples**

First of all, you should know that donating your samples is voluntary. You are under no obligation to donate your samples in order to receive medical care. Even if you agree now, if you change your mind later, you can request that any samples that have not yet been analysed are destroyed.

**Purpose of sample collection**

The samples you donate will be used for current or future studies related to COVID-19 infection. Your samples may only be used for the purpose stated in this information sheet. In the event that they may be used for other non-COVID-19 related research in the future, they must be approved by an Ethics Committee and you will be asked for a new consent form.

**Where will the tests be carried out?**

Your samples may be analysed in different laboratories of this institution, although they may be analysed in other centres, or even in foreign laboratories, depending on the types of analysis to be carried out.

**What will happen to my samples once they have been analysed?**

Your samples may only be used for the purpose stated in this information sheet. Your samples may be used in ongoing studies of COVID-19 infection, but may also be retained for future research related to COVID-19 infection.

All studies using your samples must be approved by a Research Ethics Committee. You may request to be informed about the studies in which your samples have been used.

Your samples will be stored mainly in the Germans Trias i Pujol Health Sciences Research Institute (IGTP) Biobank, an institution that has been accredited by the health authorities specifically to store samples for research under the conditions required by law. The treatment and use of all samples will be carried out in accordance with the Biomedical Research Act (14/2007) and Royal Decree 1716/2011, which regulates biobanks. The samples will be kept for 15 years. They may be transferred to third parties if the current Biobank legislation is complied with and they will not be sold under any circumstances.

In some cases, they may become part of sample collections accredited by the health authorities. In this case, they can neither be transferred to third parties nor sold, but they could later be integrated into the Biobank of our centre.

The samples will be identified in a coded form. This means that the information on the labelling of the samples cannot be linked to the person who donated the sample. Please see below the section on Confidentiality and data protection.

**Confidentiality**

The information we collect in this research project will be confidential and no one other than the researchers will be able to see it. Once all the information has been obtained, it will be transferred to a database in which an encrypted identification code will be created and no information that would allow your identity to be recognised will be recorded.

It is also important for you to know that this proposal has been reviewed and approved by the Research Ethics Committee of the IDIAP Jordi Gol and the Germans Trias i Pujol Research Institute (IGTP).

The processing of your personal data will comply with the provisions of the Organic Law 3/2018, of 5 September, on the Protection of personal data and guarantee of digital rights and Regulation (EU) 2016/679 on the protection of natural persons that relate to the processing of personal data and the free movement of such data. In accordance with the provisions of the aforementioned legislation, you may exercise the right of access, rectification, erasure, restriction of processing, portability and opposition to processing.

**Who to contact**

If you have any questions, you can ask them now or later, even after the study has started. In any case, you can contact the project managers: Pere Torán Monserrat ptoran.bnm.ics@gencat.cat, Concepció Violan Fors. cviolan@igtp.cat

If you agree to participate in this study, please express your consent by completing the document available below.