Study Protocol

Title: SARS-CoV-2 infection among health care professionals: demographic characteristics and serological and immune responses related to the phenotypic progression (ProHEPiC-19)

PIs 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.

Protocol ID: 4R20-105

Funded by the regional Ministry of Health of the Generalitat de Catalunya; (COVID19-PoC SLT16_04; NCT04885478)

December, 1 2021

INDEX

1. HEADER

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

PIs and study 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 duration: 30 May 2020 – 30 Jun 2022

Funding bodies: Departament de Salut. Generalitat de Catalunya. Call COVID19-PoC number SLT16_04, 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 characteristics, including the trajectories of symptomatology and immune response against the 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) have the most frequent contact with the virus, they are more susceptible to infection, which affects not only their own health, but it also performance of the 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 more participants.

Objectives:

Primary: To consolidate a prospective cohort of Health Care Workers (HCWs) to generate rich and dense epidemiological and clinical 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, vaccination, and long-term immune response.

Secondary:

1. To determine the kinetics of SARS-CoV-2 antibodies and cellular immune response in the early, middle, and late phases of immunization.

2. To assess the relationship between clinical variables and initial RT-PCR results considering the interindividual heterogeneity in the immune response in the early, middle, and late phases of immunization.

3. To analyse differentially expressed cytokines as biomarkers of disease trajectory in the early, middle, and late phases of immunization.

Methods and analyses: The dynamic and prospective cohort study with several follow-ups within a 12-month period has been conducted in four primary-care centres and one hospital of north metropolitan area (Metropolitana Nord) of Barcelona (Spain). At the time of writing, the study consists of 1350 participants divided into 2 cohorts: Healthy-Exposed HCWs: 436 not infected by SARS-CoV-2 (confirmed with a negative RT-PCR result and a negative test for SARS-CoV-2 antibodies at baseline) and Infected HCWs: 436 symptomatic participants (those with new persistent cough, temperature ≥37.5°C, anosmia or ageusia, or other COVID-related symptoms) and asymptomatic participants (all confirmed with a positive RT-PCR test and/or SARS-CoV-2 antibodies IgM, IgG at baseline). Primary outcomes include humoral and cellular immune response, levels of SARS-Cov-2 antibodies, symptomatology, demographics and other variables that may be predictive of the disease trajectory.

Timepoints: baseline, 15 days, 1, 3, 6, 9, 12, 15, 18, 24, 30, and 36 months post baseline

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

Research plans: This prospective cohort offers the possibility to study the relationship between the immune response kinetics and the phenotypic progression while considering the heterogeneity due to age and gender, as well as duration of the immune response. We will also be able to examine the effects

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

Ethics and dissemination: The study has been 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.

Protocol ID: 4R20-105

Keywords: SARS-CoV-2; COVID-19; cohorts; epidemiology, SARS-CoV-2 antibodies, primary health care, hospital, health care workers, immune response.

3. STUDY ABBREVIATIONS

Coronavirus Disease 2019	COVID-19
Severe acute respiratory syndrome coronavirus 2	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 the time of writing (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 Spain, where the transmission rate of SARS-CoV-2 is high, 26% of all individuals tested positive for SARS-CoV-2 by PCR were HCWs [3]. Therefore, HCWs are a high-risk group for developing the infection, that is exceptionally relevant for three reasons: 1) the clinical manifestations of the disease can be severe in some cases; 2) HCWs are caring for patients; therefore, they are not only more susceptible to infection, but they are also more likely to infect others, including co-workers, patients, and family members; and 3) the infection and the subsequent absence from work in this group reduces the capacity of the health care 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 the most suitable to be investigated for the production and kinetics of the antibodies in the early, middle and late phases of the immunisation against SARS-CoV-2

One of the major current questions is whether individuals who have been infected with SARS-CoV-2 develop long-term or even permanent immunity to the virus regardless of the symptoms. Answering this question is crucial for the epidemiologists, the policy makers, the health service professionals, and the society at large. Three different types of diagnostic tools can be used: 1) RNA-based tests using (RT-)PCR can determine whether viral genes are present in the nasal and the throat swabs, indicating active infection. 2) Presence of antibodies against SARS-CoV-2 in plasma or capillary blood can also indicate active infection and, more importantly, they also indicate whether the infected individual is able to develop a satisfactory immune response. 3) Testing the “cytokine storm” is a more exploratory approach that can include the assessment of the presence of pro inflammatory cytokines [4] in response to the SARS-CoV-2 infection. The “cytokine storm” is a particular pathogenic feature of the SARS-CoV-2 infection and has been linked to disease severity and progression [5].

Several studies on 2002–2004 SARS outbreak caused by the SARS-COV-1 have shown that levels of antibody response differ depending on the clinical spectrum [6]. Nevertheless, little is known regarding the immune response to the SARS-CoV-2. It has been observed that the IgM and the IgG against the SARS-CoV-2 antigens show the highest response, and the highest antibody responses are presented by people with a higher disease severity and admission to the hospital and the intensive care units. [7,8]. The immune response to SARS-CoV-2 infection is heterogeneous with a likely relationship with the clinical spectrum 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 immunity following an infection with the SARS-CoV-2 or that multiple exposures will be needed for a long-lasting protection. All these points are extremely important for the re-vaccination strategies.

At the current stage of the pandemic, laboratory tests measuring antibody response and seroconversion remain essential. Studies on indicators of the inflammatory response and the prognostic value of other Covid-19 indicators, such as the acute phase reactants are also needed. In addition, the identification of pro-inflammatory biomarkers may inform the disease prognosis and the immunosuppression and anti-inflammatory therapies in the acute phase. Furthermore, by conducting serological studies of the population we can : 1) better understand the kinetics of the immune response to the SARS-CoV-2; 2) determine and monitor the infection rate more accurately, which is essential to determine the probability of different disease trajectories including death; 3) identify individuals with strong antibody response,

who could be donors to convalescent serum therapies; and 4) identify who has and who lacks immunity against the disease. All of these serological information can be used to strategically deploy HCWs who are immune to the infection, thus limiting the risk of exposure and inadvertent spread of the virus.

As mentioned above, HCWs are a high-risk group for acquiring the 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 PCR tests. It should be noted that the confinement takes a minimum period of 8 days depending on the PCR result at the 8th, the 14th and the 28th days. This causes a shortage of HCWs, which seriously affects the performance of the health services and put additional pressure on the sector already overburdened by the pandemic.

The non-specificity of the initial symptoms makes it very difficult to screen for the disease. It is therefore essential to identify factors associated with the protection of professionals against SARS-CoV-2; especially because the evolution of the SARS-CoV-2 makes it likely that new epidemic waves will appear within a short period of time. To a large extent, the factors contributing to the evolution and the permanence of the virus, the immunity of the infected persons and that of healthy carriers of the infection have not been clarified.

To date, seroprevalence studies conducted in Spain have used cross-sectional designs, which cannot systematically infer on long-term antibody and cellular immune responses. Moreover, thorough serological and prospective studies are needed also to determine whether there is a pre-existing antibody-dependent enhancement (ADE) among SARS-CoV-2 infections, either because of a prior homologous infection or due to cross-reactive antibodies from other human coronaviruses causing common cold. This have a particular relevance for the vaccination plans.

Previous studies have examined various types of antibodies identified against SARS-Cov-2 (Spike1, Spike2, nucleocapside, envelopes), showing significant heterogeneity in responses [9,10]. Moreover, little is known about their simultaneous kinetics and how their responses varies according to the antigens also used in the various vaccines that will be in the market in the next months. The tracking of the fast evolution of the virus also requires an exhaustively monitored cohort that makes it possible to identify the kinetics of SARS-CoV-2-specific antibodies in individuals with different degrees of severity and to determine which levels of which antibodies are protective against possible reinfections in the longer term.

In conclusion, the present study is of utmost importance as it obtains epidemiological and clinical data to improve health policies for the prevention of and defence against the SARS-CoV-2 infection. In addition, the study will be conducted in parallel with the course of the pandemic, which allows us to obtain real-time data on the different waves that occur and, in turn, to study the SARS-CoV-2 pandemic. Finally, ProHEpIC-19 can be used as an active platform to implement SARS-CoV-2 research projects with particular emphasis on incidence rate, reinfection, vaccines, and long-term immunity.

5. HYPOTHESES AND OBJECTIVES

5.1. HYPOTHESES

The specific hypotheses are:

Infection with SARS-CoV-2 will produce an immune response with different characteristics.

1. The strength and the kinetics of the immune response depend on the clinical presentation of the COVID-19 disease.

2. The strength and the kinetics of the immune response depend on age, gender, and other clinical variables.

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

5.2. OBJECTIVES

5.2.1. PRIMARY AIM

To establish a prospective cohort of Health Care Workers (HCWs) to acquire feature-rich epidemiological and clinical data with high temporal resolution. This information will be relevant to improve health care policies and COVID-19 protocols. This cohort will also be used as an active platform to implement SARS-CoV-2 research projects with particular emphasis on incidence rate, reinfection, vaccines, and long-term immunity.

Primary Outcomes:

1. The prospective cohort of HCWs

Two groups: 436 exposed but uninfected HCW participants and 436 HCW participants infected with SARS-CoV-2

Timepoints: at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

2. Descriptive demographics (age, gender, education, housing, employment) of the cohort

Descriptive analyses including the occurrence and frequency for categorical variables, mean and standard deviation for continuous scalar variables with normal distribution, and median and inter-quartile range for continuous scalar variables with non-normal distribution10s.

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

3. Clinical description (asymptomatic, mild-moderate illness, severe-critical illness) of the cohort

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

All primary outcomes include statistical analyses comparing the two groups and the corresponding timepoints, as well as the relationships between the variables using a mixed-effects study design.

5.2.2. SECONDARY AIM

1. To determine the kinetics of SARS-CoV-2 antibodies and cellular immune response in the early, middle, and late phases of immunization.

2. To assess the relationship between clinical variables and initial RT-PCR results considering the interindividual heterogeneity in the immune response in the early, middle, and late phases of immunization.

3. To analyse differentially expressed cytokines as biomarkers of disease trajectory in the early, middle, and late phases of immunization.

6. METHODS

6.1. DESIGN

Prospective study includes two groups of health care workers (HCW) (healthy and infected). Each groups will include 436 participants.

Biological, clinical, psychological and social factors will be closely monitored at several occasions (baseline and follow-ups) during one year.

6.2. POPULATION AND SETTING

This is a multicentre study including public primary care centres in Mataró, Sabadell, and Santa Perpètua de la Moguda and the public third level hospital Hospital *Germans Tries i Pujol* located in Badalona). The HCWs are recruited from Gerència Territorial Metropolitana Nord of the Catalan Institute of Health (ICS), which consists of 7,776 HCW, including physicians, nurses, clinical researchers, medical residents, and other essential workers in direct contact with patients.

6.2.1. SELECTION CRITERIA:

HCWs at the risk of infection due to direct contact with patients. Different recruitment strategies are applied to contact with participants, such as email and phone calls.

Inclusion criteria:

≥ 18 years of age

Accepting to take part in the study and providing the informed consent

HCWs potentially exposed (HCW that works in the same place as the HCW infected) to SARS-CoV-2.

Exclusion criteria:

Refusing taking part in the study and/or providing the informed consent

HCW unlikely to be exposed (HCW that works in the same place as the HCW infected) to SARS-CoV-2

6.3. STUDY PERIOD

From 31st March, 2020 to 15th June 2021 (estimated).

6.4. SAMPLE SIZE (POWER)

As of 23rd March 2020, 1350 people meet eligibility criteria.

We include 436 participants in each group and investigate the course of immunity to the SARS-CoV-2. The sample will be representative of the study population.

Sample Size Estimation

Assuming a Test for Two Means in a Repeated Measures Design, group sample sizes of 436 and 436 achieve 90% power to detect a difference of 0.2 in a design with 13 repeated measurements having a AR(1) covariance structure when the standard deviation is 2, the correlation between observations on the same subject is 0.5, and the alpha level is 0.05 [11-13].

Dropout-Inflated Sample Size

Groups	Dropout Rate	Sample Size	Inflated Enrollment Size	Expected Number of Dropouts
1 and 2	10%	436	485	49
Total		872	970	98

Based on these estimates, our planned sample size of 436 per group ensures sufficient power to detect the effects.

6.5. PROCEDURES

Participants entering the study after meeting inclusion/exclusion criteria (Table 1) give written informed consent (for informed consent form and patient information sheet, see Appendix 1), and they are assigned to a group (Figure 1):

1) Healthy group: HCW exposed to but not infected with SARS-CoV-2, negative test for virus RNA (PCR) or IgM and IgG against the nucleocapsid (N) of the virus

2) HCW infected with SARS-CoV-2 based on RT-PCR, as well as IgM(N) and IgG(N) antibody tests

All participants undergo a standardised protocol according to the specific group to which they are assigned to (Table 2). At study entry, we record the following information: socio-demographic data, clinical data (symptoms), epidemiological data (contact with COVID-19 positive person(s) 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, the initial group assignments are reviewed and finalised based on laboratory results (RT-PCR test and SARS-CoV-2 antibodies IgM(N) and IgG(N):

1) Healthy HCWs (HCW_H): This group includes HCWs with negative laboratory tests for virus RNA and SARS-CoV-2 antibodies IgM(N), and IgG(N).

2) Infected HCWs (HCW_I): This group includes HCWs with positive laboratory tests for virus RNA and SARS-CoV-2 antibodies IgM(N) and IgG(N). This group is subdivided into three subgroups according to illness severity:

Asymptomatic or Presymptomatic Infection: Individuals with positive laboratory evidence of COVID-19 but no COVID-19 signs or symptoms.

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

Severe-Critical Illness: Individuals with positive laboratory evidence of COVID-19 who have any of the typical symptoms of COVID-19 (e.g., fever, cough, sore throat, malaise, headache, muscle pain, nausea, vomiting, diarrhoea, loss of taste and smell), also have shortness of breath, dyspnoea or positive chest imaging, and treated in the hospital.

The whole assignment process is conducted by a trained multidisciplinary research staff, including family practice physicians, nurses, assistant manager, and project manager (Figure 1).

6.5.1. RECRUITMENT

The Basic Prevention Unit of the Northern Metropolitan Area of Barcelona in Spain (NM_BCN) offers HCW to participate in the ProHEPIC-19 study. If they accept, then the researchers of the ProHEpiC-19 contact these prospective participants.

Eligible participants interested in the study, are exhaustively informed about the ProHEpiC-19 protocol and provide an informed consent to participate.

Participants complete several clinical questionnaires and undergo an immunological assessment : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.. (Table 2)

Participant data will be anonymised with a numerical code system and stored securely as detailed in section 12 (Data storage and security). . Data storage will be maintained 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 provide an external validity of the recruited sample of participants.

6.6. VARIABLES AND MAIN COVARIABLES

1) Gender measured as a categorical variable with two categories (Male, Female)

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

3) Education: measured as an ordinal variable with 5 levels (has not completed primary education, completed primary education, completed secondary education, completed post-secondary non-tertiary education, completed short-cycle tertiary education, completed higher education (university)).

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

5) Employment (professional category, schedule, workplace characteristics)

6) Clinical presentation of COVID-19: defined according to national and international guidelines (asymptomatic, mild-moderate , severe-critical) as defined in section 6.5

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

8) Date of onset: the date of appearance of the first symptom(s) as described in Table 2 and further confirmed by a positive RT-PCR test or antigen (IgM(N) or IgG(N)) test.

9) Diagnostic code: according to the International Classification of Diseases, version 10 (ICD-10), including the specific codes for COVID-19.

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

11) Medication: classification number of drugs according to the Anatomical Therapeutic Chemical (ATC) classification system.

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

13) Comorbidities: diabetes, hypertension, ischaemic heart diseases, other chronic comorbidities (ICD-10 codes)

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

15) Psychosocial and health-related habits: (sleeping, eating, sexual activity, physical activity, and family relationship)

16) Laboratory Assays:

A) RT-PCR will be performed in every 2 weeks. RNA will be extracted from fresh samples using the STARMag 2019-nCoV kit (Seegene, CA, USA) and 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 a 10' of incubation at room temperature, 350 µl will be used for nucleic acid extraction resulting in a volume of 100 µl of extracted genome, from which 8 µl will be taken for the 2019-nCoV PCR assay. The detection of RNA will be performed using the Allplex™ SARS-CoV-2 Assay, which is a multiplex real-time PCR assay to detect 4 target genes of the SARS-CoV-2 in a single tube. According to the manufacturer's instructions, the assay is designed to detect the RdRP, S, and N genes specific for the SARS-CoV-2, and the E gene specific for all Sarbecovirus including the SARS-CoV-2 (Seegene, CA, USA).

B) Quantification of antibodies to SARS-CoV-2: We will perform the immunological survey : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study. We performed a previous validations study with six different IVD-CE approved and commercially available ELISA test kits and selected the AntiSARS-CoV-2 IgG and IgM ELISA (Inmunodiagnostic). This test 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 a test was considered positive with a test value >1.1, undetermined with a test value between 0.9 and 1.1 and negative with a test value <0.9 [14]

Infected participants will be also tested for antibodies against the Spike (S) subunit of SARS-CoV-2using the DECOV1901 enzyme immunoassay. 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 a test was considered positive with a test value >40, undetermined with a test value between 32 and 40, and negative with a test value <32 Ul/ml [15].

C) Additional Immune response laboratory assays:

C1) INF-γ Elispot assays to monitor SARS-CoV-2 specific CD4+ and CD8+ T-cell responses Cryopreserved peripheral blood mononuclear cells (PBMCs) will be thawed and incubated overnight at 37°C in the presence of a SARS-CoV2 peptide pool (5ug/ml for each peptide) and phytohemagglutinin (PHA) (15 µg/ml), as positive control. Finally, plates will be treated with biotinylated anti-human IFN-γ, streptavidin-alkaline phosphatase and its coloured substrate (Mabtech, Sweeden). IFN-γ secreting cells will be quantified using Immuno Capture and Immuno Spot software to calculate the number of Spot Forming Cells (SFC). Test will be considered positive if it containes at least 50 SFC per 106 PBMCs above the background level (2x mean + 3x standard deviation).

C2) 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). Test will be repeated for confirmation.

C3) Combinatorial flow-cytometry of lineage, activation markers and intracellular cytokines: Cryopreserved PBMCs will be stimulated with reactive peptide pools identified by ELISpot and incubated overnight. 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).

C4) Collected blood samples will be stored for more exhaustive determinations of the immune response in later studies.

17) Vaccines, we registered all possible COVID 19 vaccinations (Astra- Zeneca, Pzifer_ Biontech; J&J_Janssen; Sanofi_GSK; Moderna_Lonza; Curevac; Novavax) Date 1st. 2nd., 3rd, dose.

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 to investigate possible relationships between possible prognostic factors and the immune response.

Primary Outcome Measure:

1. The prospective cohort of HCWs

Two groups: 436 exposed but uninfected HCW participants and 436 HCW participants infected with SARS-CoV-2.

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

2. Descriptive demographics (age, gender, education, housing, employment) of the cohort

Descriptive analyses including the occurrence and frequency for categorical variables, mean and standard deviation for continuous scalar variables with normal distribution, and median and inter-quartile range for continuous scalar variables with non-normal distributions.

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

3. Clinical description (asymptomatic, mild-moderate illness, severe-critical illness) of the cohort

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

All primary outcomes include statistical analyses comparing the two groups and the corresponding timepoints, as well as the relationships between the variables using a mixed-effects study design.

Secondary Outcome Measures:

4. Kinetics of SARS-CoV-2-specific IgM(N)

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

5. Kinetics of SARS-CoV-2-specific IgG(N)

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

6. Kinetics of SARS-CoV-2-specific IgG(S)

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

7. Kinetics of T-Cell

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

8. Assessment of the relationship between clinical variables and the initial RT-PCR results in the sample as a whole and stratified by gender: We will use mixed-effect ANOVAs or Kruskal-Wallis followed by Holm-adjusted Dunn's test, after checking normality assumption using a Shapiro-test.

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

9. Assessment of the relationship between clinical variables and the interindividual heterogeneity in the immune response in the early, middle, and long late phases of immunization in the sample as a whole and stratified by gender: We will use mixed-effect ANOVAs or Kruskal-Wallis followed by Holm-adjusted

Dunn's test, after checking normality assumption using a Shapiro-test. Similarly, to look for differences in antibody levels between males and females, either a t-test or a Mann-Whitney test will be performed.

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

10. Cytokines as biomarkers of disease progression in the early, middle, and long late phases of immunization: Appropriate statistical tests (i.e. t-test or Mann-Whitney to compare between males and females and ANOVA or Kruskal-Wallis followed by Holm-adjusted Dunn's test) to compare between clinical presentations) will be used after checking for normality (Shapiro-test).

Timepoints: : at baseline, 7, 15 days and 1,2 3, 6, 9 12, 15, 18, 24, 30, and 36 months after the beginning of the study.

6.7. Data sources

We will use various information sources.

A) ProHEpiC-19 cohort.

The data will be collected and stored in an internal centralized data warehouse at the Catalan Health Institute. After data collection and curation, the dataset will be (1) validated for the accuracy of demographic, clinical, social (2) extended with the prospective acquisition of laboratory tests.

B) Primary care electronic health records (e-Cap) containing demographic data and clinical variables: health conditions (ICD-10 codes) and diagnostic data, medication, comorbidities, laboratory test

C) Study-specific Case Report Form (Teleform) including a unique identifier for the patient

6.8. STATISTICAL ANALYSIS

Data collection: The data will be reformatted for the purpose of the analysis and to meet the specific requirements of the statistical packages used in the study. A comprehensive data scrubbing and validation will be conducted to eliminate duplicities, missing data, and treatment indications inconsistent with the diagnosis. Results will be accepted as statistically significant if their statistical probability will be at or lower than 0.05. All the analyses will be performed using R version 3.6.3 or higher.

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

Descriptive analyses will be performed to characterise the cohort at each timepoints with the occurrence and frequency for categorical variables, mean and standard deviation for continuous scalar variables with normal distribution, and median and inter-quartile range for continuous scalar variables with non-normal distributions. Frequency and percentage of dropouts will be updated at each timepoint. Characteristics of the cohort and the dropouts will be assessed using appropriate statistical tests in order to prevent potential biases due to dropouts. Groups within the cohort will be compared at each time point in terms of sociodemographic, epidemiological, clinical, and immunological information available using parametric and non-parametric tests depending on the number and the nature of the included variables tested with Shapiro-test. Group comparisons at singe timepoints will employ t-test or Mann-Whitney u-test (two groups) and ANOVA or Kruskal-Wallis test (three groups).

6.8.2. SECONDARY AIMS

6.8.2.1. TO DETERMINE THE KINETICS OF SARS-COV-2 ANTIBODIES AND CELLULAR IMMUNE RESPONSE IN THE EARLY, MIDDLE, AND LATE PHASES OF IMMUNIZATION.

Two techniques will be employed to study the kinetics of SARS-CoV-2 antibodies since symptoms onset. Locally Estimated Scatterplot Smoothing (LOESS) models will be fitted as a descriptive approach. Then, nonlinear mixed-effects (NLME) models will be fitted to obtain the curves characterising the kinetics.

LOCALLY ESTIMATED SCATTERPLOT SMOOTHING (LOESS) MODELS

LOESS models will fit to describe the evolution of the antibodies, both IgM and IgG, since the diagnosis. We will use data only from participants with positive tests. LOESS regression consists of fitting first or second-degree polynomials in subsets of data determined by a nearest-neighbours algorithm. The resulted model parameters will be stratified according to clinical condition and gender and will be described with their mean and 95% confidence interval (CI).

NONLINEAR MIXED-EFFECTS (NLME) MODELS

Mixed-effects models are generally useful when there are multiple measures per participant, as in this study. Random effects allow to generalise the results from a random sample to the whole population from which the sample has been drawn. It compares the variance explained by the modelled effect(s) to that of that due to the random sampling. The nonlinear models typically more flexible and able to describe more complex patterns in the data than their linear counterparts. However, they also require more data to avoid overfitting. Therefore, nonlinear models provide the best approach to characterise our cohort.

We will use NLME models to investigate the kinetics of SARS-CoV-2 antibody (IgM, IgG) levels since the onset, as well as how the kinetics varies according to clinical presentation and gender. We will use data only from participants with positive tests. Time after diagnosis will be discretized as follows: Tests made in the first 14 days since diagnosis are treated as "Day 0". Tests from days 15 to 29, as "Day 15", days 30 to 59 as "Day 30", days 60 to 89 as "Day 60", days 90 to 179 as "Day 90", days 180 to 269 as "Day 180", days 270 to 360 as "Day 270", days 360 to 450 as “Day 360”, and days 450 up to end of follow-up as “540”.

The general equations to be fitted will be:

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

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

t has been 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 specified and estimated using the nlme package [16].

6.8.2.2. TO ASSESS THE RELATIONSHIP BETWEEN CLINICAL VARIABLES AND INITIAL RT-PCR RESULTS CONSIDERING THE INTERINDIVIDUAL HETEROGENEITY IN THE IMMUNE RESPONSE IN THE EARLY, MIDDLE, AND LATE PHASES OF IMMUNIZATION

The relationship between the incidence of SARS-CoV-2 and sociodemographic, epidemiological, clinical, and immunological factors will be investigated using parametric and non-parametric tests depending on the number and the nature of the included variables tested with Shapiro-test. Group comparisons at singe timepoints will employ t-test or Mann-Whitney u-test (two groups) and ANOVA or Kruskal-Wallis test (three groups)., while analysis of interactions between variables at single timepoints will employ Chi Square test or ANOVA.

6.8.2.3. TO ANALYZE DIFFERENTIALLY EXPRESSED CYTOKINES AS BIOMARKERS OF DISEASE PROGRESSION IN THE EARLY, MIDDLE, AND LATE PHASES OF IMMUNIZATION

Cytokine levels will be studied and compared across different clinical presentations and gender to validate their feasibility as biomarkers 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).

6.9 ETHICS

The biological samples for this investigation will be available and stored according to different study protocols approved by the Institutional Review Board. Study participants will be informed about the objectives of the study and the interventions linked to their participation. They will receive a written information sheet detailing all the information about the project and asked to provide a written informed consent. The confidentiality and anonymity of the data will be ensured in accordance with current regulations (Organic Law 3/2018, of 5 December, on Personal Data Protection and guarantee of digital rights). The study will be conducted in accordance with the articles pertaining to the Declaration of Helsinki, agreed at the 64th General Assembly in 2013. The project is under the tutelage of the Clinical Research and Ethics Committee (CEIC) of the Instituto Universitario de Investigación en Atención Primaria (IDIAPJGol ref 20/067-PCV IGTP ref COV20/00660 (PI-20-205)) and incorporates its recommendations and suggestions.

The samples may only be used for the purpose stated in the information sheet and may be used in ongoing and future studies of COVID-19 pandemic.

All studies using the samples must be approved by a Clinical Research and Ethics Committee, and the participants may request to be informed about the studies in which the samples will have been used.

The samples will be stored mainly at the Germans Trias i Pujol Health Sciences Research Institute (IGTP) Biobank, that has been accredited by the medical regulation authorities *(Insitituto de Salud Carlos III. Ministerio de Ciencia e Innnovación*) for the specific purpose to store samples for research under the conditions required by law (number B.0000643) and the 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 according to the current legislation , but they will never be sold under any circumstances.

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

The samples will be anonymised according to a numeric coding system. This means that the data cannot be linked to the participant.

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 Fajgenbaum DC, June CH. Cytokine Storm. N Engl J Med. 2020 Dec 3;383(23):2255-2273. doi: 10.1056/NEJMra2026131. PMID: 33264547; PMCID: PMC7727315.

5. Metha P, McAuley DF, Brown M, Sannchez E, Tattersall RS, Manson JJ.COVID- 19:consider cytokine storm syndromes and immunosuppression. The Lancet 2020. doi.org/10.1016/S0140-6736(20)30628-0

6. 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.

7. 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

8. 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.

9. 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).

10. 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.

11. Brown, H. and Prescott, R., 2006. Applied Mixed Models in Medicine. 2nd ed. John Wiley & Sons Ltd. Chichester, West Sussex, England. Chapter 6.

12. Liu, H. and Wu, T., 2005. 'Sample Size Calculation and Power Analysis of Time-Averaged Difference.' Journal of Modern Applied Statistical Methods, Vol. 4, No. 2, pages 434-445.

13. Diggle, P.J., Liang, K.Y., and Zeger, S.L., 1994. Analysis of Longitudinal Data. Oxford University Press. New York, New York. Chapter 2.

14. IgG N: https://www.immunodiagnostics.com.hk/product-page/sars-cov-2-np-igg-elisa-kit-ce-ivd

IgM N: https://www.immunodiagnostics.com.hk/product-page/sars-cov-2-np-igm-elisa-kit-ce-ivd

15. IgG S: https://www.demeditec.com/en/products/covid-19-sars-cov-2-igg-elisa-decov1901/ifu-decov1901-covid-19-sars-cov-2-igg-elisa-210401.pdf

16. 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 (sorted according to their affiliations):

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

9. MODIFICATION OF THE PROTOCOL

Changes from the original plan will be summarized on a separate document, and new version of study protocol will be created.

10. DISSEMINATION

To increased awareness and maximise the impact of our project, we put special emphasis on the improvement of communication at every level and, in particular between health care professionals and patients.

In support of the project's objectives, a communications strategy will be developed considering the various potential target groups to provide them with consistent overall messages and ensure their long-term engagement. The strategy will include:

a) Project visual identity (websites, poster, advertisements)

b) Key messages

c) Timing and content of research briefings and policy notes

d) Schedule of public activities.

At the same time, social media channels for the project will be established. The strategy will detail dissemination of specific outputs to the public media. The activities detailed in the strategy will be reviewed every six months.

Communication activities

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

a) Social networks: Engage with over 100 stakeholders by the end of 12, 24, 36 months , and 1000 by the end of the project.

b) Briefing notes: At least six short, briefing notes produced in layman terms and distributed

c) Press Releases: At least three briefing notes produced in layman terms and distributed

d) Publications and conference presentations: At least three scientific manuscripts and six conference presentations by the end of the project including a final conference. Publications and conference presentations will be targeted and timed according to the findings and to utilise other professional conference and journal special issues.

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 pseudo-anonymised by coding the identifying variables in both the initial paper-based data collection booklet (CB) and the subsequently created database. Only data necessary to achieve the research objectives will be collected. The necessary variables are: 1) 1) Biological gender; 2) Age (years); 3) Education; 4) Housing; 5) Employement; 6) Clinical presentation; 7) Clinical monitoring of COVID-19 symptoms; 8) Date of onset; 9) Diagnostic code; 10) Use of personal protective equipment; 11) Medication; 12) Hospital admissions; 13) Comorbidities; 14) Mortality by any causes; 15) Psychosocial and health-related habits; 16) Laboratory Assays: RT-.PCR test, ELISA for IgM and IgG against the nucleocapside protein, IgG against the spike protein) , INF-γ Elispot assays; Citokines multiplex assay; Combinatorial flow-cytometry of lineage, activation markers and intracellular cytokines.

In accordance with regulation 14/2007, the biological samples of each participant in the project will be collected directly from the participant and in accordance with the procedure set out in the regulations on the use of biological samples.

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 and 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.

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

The necessary variables will be obtained directly from the participants by means of a self-completed form. 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 GDPR, 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 the data controllers for this study. The project database will be stored in the computers of the Unitat de Suport a la Recerca Metropolitana Nord of the ICS, which will act as the data processor. International data transfer is not expected at this stage. 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 will be created and stored on a physical computer where the organisation has installed the corresponding software and therefore only accessible from computers with 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 regulations.

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, the companies of Closca, Leroy Merlin, Barfim, and Roche for providing us with personal protection equipment and medical materials.

In addition, the researchers would like to thank Núria Prat Gil and Josep Ma Bonet Simó, Director and 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. The Informed Consent Form and Participant Information Sheet are also- available in English.

All written information and other material to be used by the participants and the staff must phrased in layman terms.

16. COMPENSATION

Participants do not receive any compensation in any kind.

17. OPEN ACCESS

A public dissemination of the results will be made in indexed journals in different subjects categories in 1st quartile

18. PATIENT CONSENT DETAIL

The Informed Consent Form and the Patient Information Sheet are attached to this protocol in the Appendices section.

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 any researcher or institution outside the project is 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. APPENDICES

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

Table 1. ProHEpiC-19 Selection criteria.

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

Table 2 Participant assignment procedure*

Visit	Pre-cohorts enrolled		Protocol
0	**All possible participants**		Telephone screening/personal interview to explain the study and to decide if they are eligible to participate
1	**Healthy exposed healthcare workers**		RT-PCR (SARS-CoV2) Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S) Hemogram and serumbiochemistry
	Recently infected healthcare workers		RT-PCR (SARS-CoV2) Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S) Hemogram and serum biochemistry T-cells and cytokines
	Previously infected healthcare workers		No visit
2	**Healthy exposed healthcare workers and**		No visit Symptom control call via phone
	Recent infected healthcare workers		Symptom control call via phone
	No-recent infected healthcare workers		No visit
3	**Healthy exposed healthcare workers and**		No visit
	Recently infected healthcare workers	With positive RT-PCR in Visit 1 or previously	RT-PCR (SARS-CoV2) Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)
		With negative RT-PCR in Visit 1 or previously	Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
	Previously infected healthcare workers		No visit
4	**Healthy exposed healthcare workers**		Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)

<table>
<tr><td></td><td colspan="3"></td><td></td></tr>
<tr><td></td><td colspan="3">Recently infected healthcare workers</td><td>Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)
T-cells and cytokines</td></tr>
<tr><td></td><td colspan="3">Previously infected healthcare workers</td><td>RT-PCR (SARS-CoV2)
Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)
Hemogram and serum biochemistry</td></tr>
<tr><td rowspan="8">5</td><td rowspan="2">Healthy exposed healthcare workers</td><td colspan="2">With antibodies in Visit 4</td><td>Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)
T-cells and cytokines</td></tr>
<tr><td colspan="2">Without antibodies in Visit 4</td><td>Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)</td></tr>
<tr><td rowspan="3">Recently infected healthcare workers</td><td colspan="2">With positive RT-PCR in Visit 1 or previously</td><td>Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)
T-cells and cytokines</td></tr>
<tr><td rowspan="2">With negative RT-PCR in Visit 1 or previously</td><td>With antibodies in Visit 4</td><td>Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)T-cells and cytokines</td></tr>
<tr><td>Without antibodies in Visit 4</td><td>Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)</td></tr>
<tr><td rowspan="2">Previously infected healthcare workers</td><td colspan="2">With antibodies in Visit 4</td><td>Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)
T-cells and cytokines</td></tr>
<tr><td colspan="2">Without antibodies in Visit 4</td><td>Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)</td></tr>
</table>

6	**Healthy exposed healthcare workers**	With antibodies in Visit 4		Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S) T-cells and cytokines
		Without antibodies in Visit 4		Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
	Recently infected healthcare workers	With positive RT-PCR in Visit 1 or previously		Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S) T-cells and cytokines
		With negative RT-PCR in Visit 1 or previously	With antibodies in Visit 4	Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) T-cells and cytokines
			Without antibodies in Visit 4	Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
	Previously infected healthcare workers	With antibodies in Visit 4		Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S) T-cells and cytokines
		Without antibodies in Visit 4		Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
7	**Healthy exposed healthcare workers**			Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)
	Recently infected healthcare workers			Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S) T-cells and cytokines
	Previously infected healthcare workers			Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)

8	**Healthy exposed healthcare workers**	Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N)
	Recently infected healthcare workers with positive RT-PCR in Visit 1 or previously	Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S) T-cells and cytokines
	Previously infected healthcare workers	Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)
9	**Healthy exposed healthcare workers**	Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S)
	Recently infected healthcare workers	Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S) T-cells and cytokines
	Previously infected healthcare workers	Anti-SARS-CoV-2 IgG and IgM antibodies against the nucleopcapside protein, (N) Anti-SARS-CoV-2 IgG antibodies against the Spike protein, (S))

* Since visit 10 (15 month of follow-up) we wiil do the same protocol as 9 visit

Table 3 . Initial symptoms of COVID-19

SARS-COV2 CLINICAL SYMPTOMATOLOGY
Headache
Diarrhea
Dyspnea
Fever
Cough
Anosmia
Arthralgia~~s~~
Asthenia
Shivers
Thoracic pain
Epigastralgia
Discomfort
Myalgias
Nausea
Odynophagie
Congestion

Table 4 Symptom monitoring*

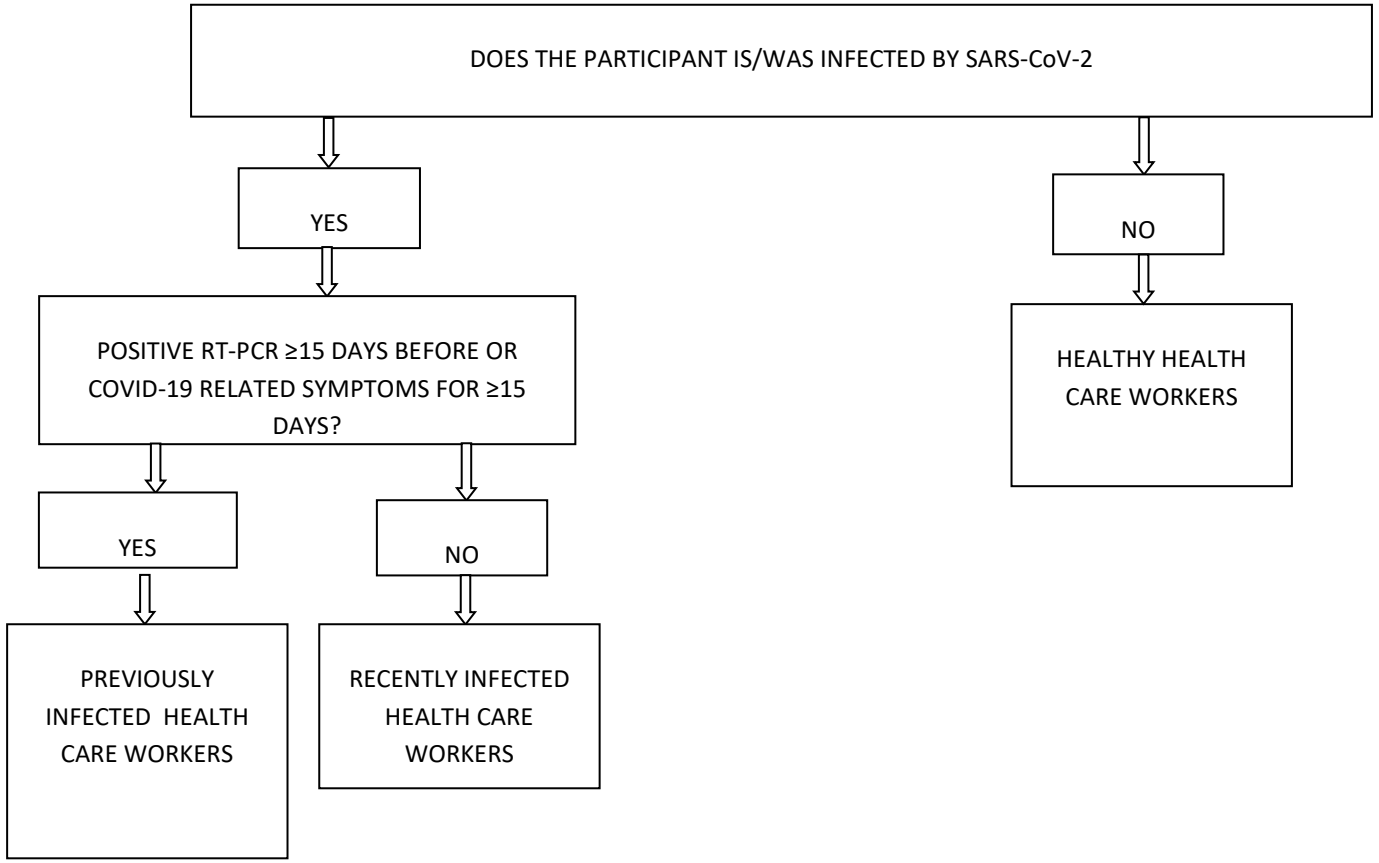
Grouping Symptom name	Date of onset	Duration (days)	Notes
Cardiovascular			
Fainting			
Pain/burning in chest			
Tachycardia			
Bradycardia			
Palpitations			
Visibly inflamed/ bulging veins			
Dermatologic			
COVID toe			
Dermatographia			
Other skin and allergy			
Peeling skin			
Petechiae			
Skin 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			
Dizziness, 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 (with normal O_2 saturation level)			
Cough with mucus production			
Coughing blood			
Other respiratory and sinus			
Shortness of breath			
Sneezing			
Reproductive/genitourinary/endocrine			
Any menstrual issue			
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*

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 phenotypic progression.**

I, ...

I have read the attached participant information sheet.

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 my data.

I consent voluntarily to participate 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 phenotypic progression**

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

This sheet provides information about this research and invite you 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.

Research objectives

At the moment, healthcare workers 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; however, the factors influencing the development of the infection, the length of the virus presence and the length of the immunity after infection are largely unknown, and there is no clear procedure to identify healthy carriers. The main objective of this research is to establish a cohort of health professionals exposed to SARS-CoV-2 to capture 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 medical institutes.

Type of research and procedures

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

1. At least two swab samples will be taken from oropharyngeal and nasopharyngeal exudates for the RT-PCR diagnostic test to detect the presence of the virus. The first sample will be taken upon the first visit, the second sample 14 days later, 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 RT-PCR test to measure C-reactive protein in the blood as an inflammation marker.
3. Six blood samples obtained by venous puncture will be taken for serological and immunological tests. The first one upon the first visit; the second, third and fourth 14, 60, 90 days later; the fifth between days 180 and 270; and the sixth at the end of the study, i.e. one year after the first visit.

4. A questionnaire on demographic, clinical, social, and psychological information will be administered during the first month after the first visit. 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 informed consent, validated by the Research Ethics Committee (REC).

Selection of participants

We are inviting 600 healthcare workers already involved in the assistance to people in isolation due to SARS-CoV-2 exposure and a further 300 healthcare workers with different risks of infection, all systematically selected from the Northern Metropolitan Territorial Management of the Catalan Health Institute (ICS).

Voluntary participation

Your participation in this research is completely voluntary.

Benefits

The results of this study will contribute to the characterisation and identification of the carrier status and immune response in healthcare workers with various levels of exposure to SARS-CoV-2. This will optimise the management of human resources with clear criteria for the safety of the professionals and the patients, thus supporting the healthcare system to cope with this pandemic and potential new waves. 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, 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, you can change your mind later and request that any samples that have not yet been analysed are destroyed.

Purpose of sample collection

The samples you donate may be used for future studies related to COVID-19 infection in addition to the purpose stated in this information sheet. In case they may be used for other non-COVID-19 related research in the future, a new approval from an Ethics Committee and a new informed consent from you must be obtained.

Where will the tests be carried out?

Your samples may be analysed in different laboratories of this institution, and 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 and 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 our biobank.

The samples will be anonymised according to a numeric coding system. 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 random numeric identification code will be created and no information that would allow your identification 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, later, or even after the study has started. In any case, you can contact the project managers: Pere Torán Monserrat ptoran.bnm.ics@gencat.cat and Concepció Violan Fors. cviolan@igtp.cat

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