Connecting European Cohorts
to Increase Common
and Effective Response to
SARS- CoV-2 Pandemic

DELIVERABLE

WP8 D8.3

Risk associated with socio-economic and environmental factors

Barcelona Institute of Global Health (ISGlobal)

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

Project Classification

Project Acronym:	ORCHESTRA
Project Title:	Connecting European Cohorts to Increase Common and Effective Response to SARS- CoV-2 Pandemic
Coordinator:	UNIVR
Grant Agreement Number:	101016167
Funding Scheme:	Horizon 2020
Start:	1st December 2020
Duration:	36 months
Website:	www.orchestra-cohort.eu
Email:	info@orchestra.eu

Document Classification

WP No:	WP8
Deliverable No:	D8.3
Title:	Risk associated with socio-economic and environmental factors
Lead Beneficiary:	ISGlobal
Other Involved Beneficiaries:	
Nature:	Report
Dissemination Level:	Public
Due Delivery Date:	M30
Submission Date:	M30
Justification of delay:	N/A
Status:	Final
Version:	1.0
Author(s):	Elisa Sicuri, Iris Lopes-Rafegas

History of Changes

Version	Date	Created/Modified by
1.0	06/06/2023	Iris Lopes-Rafegas, Elisa Sicuri

ORCHESTRA has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101016167

Table of contents

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

Executive summary

<u>Content</u>: This document includes an analysis of demographic and socioeconomic risk factors associated with Hospital and ICU admission during the acute phase and post-COVID-19 LTS, using a prospective cohort of SARS-CoV-2 infected patients from WP2 including patients from 5 countries (Argentina, Italy, Spain, Netherlands, and France). In addition, we report on the work published by Kogevinas et al (2021) and Kogevinas et al (2023), on *"Ambient Air Pollution in Relation to SARS-CoV-2 Infection, Antibody Response, and COVID-19 Disease"* and *"Long-Term Exposure to Air Pollution and COVID-19 Vaccine Antibody Response"*, using a Cohort Study in Catalonia, Spain (COVICAT Study).

Demographic and socio-economic risk factors of Hospital and ICU admission and post-COVID-19 LTS

Until 7th October 2022, 3668 patients completed the 12-month follow-up. Of those patients, 1796 had information on the post-COVID-19 LTS outcome, from which 1030 (57%) presented at least one symptom after 12 months from acute infection. 2879 (78%) patients were hospitalised and 885 (24%) had an ICU admission. After controlling for clinical risk factors and cohort fixed effects, no demographic or socio-economic factors were associated with post-COVID-19 LTS, except female sex. Several factors were associated with an increased likelihood of hospital admission for COVID-19: male sex, older age, infection during the first wave of the pandemic, and informal employment. Students and healthcare workers were less likely to be admitted to the hospital. Only male sex was significantly associated with ICU admission ($p<0.001$). Fragile populations (positive) and healthcare workers (negative) were weakly associated with ICU admission.

Environmental risk factors of SARS-CoV-2 Infection, Antibody Response, COVID-19 Disease, and COVID-19 Vaccine Antibody Response

Among those tested for SARS-CoV-2 antibodies, 743 (18.1%) were seropositive and among all participants, 481 (5.0%) had COVID-19 disease. Among infected participants, exposure to NO2 and PM2.5 was positively associated with IgG levels for all viral target antigens. Air pollution levels were also associated with COVID-19 disease. Associations of air pollution with COVID-19 disease were more pronounced for severe COVID-19. Air pollution levels were not statistically significantly associated with SARS-CoV-2 infection. Among vaccinated persons not infected by SARS-CoV-2 (n=632), higher pre-pandemic air pollution levels were associated with lower vaccine antibody response for IgM (one-month post-vaccination) and IgG. Differences in IgG levels by air pollution levels persisted with time since vaccination. No association of air pollution with vaccine antibody response among participants with prior infection (n=295) was observed.

<u>Dissemination level:</u> *This document is public.*

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

Core content

General Introduction

The highly contagious nature of SARS-CoV-2 implied a rapid spread worldwide leading to significant morbidity and mortality, straining healthcare systems, and causing disruptions to socio-economic systems. The analysis of socio-economic and environmental risk factors associated with SARS-CoV-2 infection, hospitalization, and post-COVID-19 long-term sequelae (LTS) is of importance in understanding the influence of SARS-CoV-2 on individuals and communities from a wider perspective and can help identify vulnerable populations who may be disproportionately affected by the virus: low socio-economic status has been assessed as a risk factor during the first wave of the pandemic [1]. On one hand, vulnerable populations are more likely to have poor health, increasing the risk of infection and (severe) disease [2]. On the other hand, people working in poor conditions, subject to crowded public transportation means, lacking the opportunity to work from home, or living in overcrowded accommodations, among other factors, faced increased exposure to the infection [3]. Beyond these general considerations, socio-economic gradients in COVID-19 have shown considerable variation across time and geography. Socio-economic gradients in the second and third waves were different from those observed in the first one, and such variation was not equal across regions and countries. For example, in Germany, the social stratification of COVID-19 changed substantially across the two phases of the pandemic and only in the second phase COVID-19 predominantly hit the poorest districts [4]. Therefore, factors such as income level, education, occupation, and access to healthcare services play crucial roles in determining individuals' susceptibility to infection, severity of illness, and various other clinical outcomes. By evaluating these risk factors, we can gain insights into the underlying disparities and inequities that influenced the spread and health repercussion of the virus.

Beyond the acute phase, studying the post-COVID-19 long-term sequelae (LTS) is of relevance for comprehensive patient care guidelines and resource allocation. Some individuals experience persistent symptoms and complications long after the acute infection, commonly referred to as “long COVID-19” or “post-acute sequelae of COVID-19 (PASC)” (amongst other names). Beyond the already analysed clinical factors (WP2 D2.6), the analysis of socio-economic factors associated with post-COVID-19 LTS can provide a deeper understanding of the full dimensionality of the disease and help develop better strategies to support affected individuals.

We use the prospective cohort of hospitalised patients from WP2 to extend the analysis of clinical risk factors reported in WP2 D.2.6. Beyond age, gender, and clinical factors, we study the influence of socio-economic factors, such as ethnicity, education, labour situation and healthcare worker status, into the presentation of post-COVID-19 LTS. The different factors' influence will be determined using statistical models such as multivariate regression. In addition, we evaluate the association of the same socio-economic factors with two outcomes in the acute phase: hospital admission and ICU admission.

Furthermore, we report on the work published by Kogevinas et al (2021) and Kogevinas et al (2023) on *“Ambient Air Pollution in Relation to SARS-CoV-2 Infection, Antibody Response, and COVID-19 Disease”* and *“Long-Term Exposure to Air Pollution and COVID-19 Vaccine Antibody Response”*, respectively, using a Cohort Study in Catalonia, Spain (COVICAT Study). Environmental factors, such as air pollution, can significantly influence the likelihood of viral transmission and disease severity. Air pollution had been initially associated with COVID-19 disease in ecological studies and later in cohort studies using individual data [5-10]. There are

ORCHESTRA has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101016167

some differences between the findings of individual-based studies concerning specific pollutants[11] and association with clinical disease [12], but overall results are consistent in showing that long-term exposure to air pollution is associated with COVID-19 disease and severity of the disease. Even though such studies are subject to potential biases, particularly selection bias and confounding [13], most recent evidence indicates a positive association between long-term exposure to air pollution and COVID-19 hospitalizations and severity. Biases identified in early studies may still be present in more recent studies, but they are likely to be better addressed. In addition, air pollutants have the potential to affect immune pathways involved in COVID-19 and may influence vaccine efficacy. However, studies on the effects of air pollution on post-vaccination antibody levels were limited at the time of the study reported [15].

In the above-mentioned studies, the authors show that pre-pandemic exposure to air pollution in Catalonia was associated with a 20%-50% increased risk of COVID-19 disease and a higher risk for severe COVID-19 [14]; and that higher pre-pandemic air pollution levels were associated with lower vaccine antibody response for IgM (one-month post-vaccination) and IgG, among vaccinated persons not infected by SARS-CoV-2 (n=632) [15]. These studies were conducted prior to the COVICAT cohort joining ORCHESTRA-WP3 and, thus, the study was not funded by ORCHESTRA.

Overall, the analysis of socio-economic and environmental risk factors associated with SARS-CoV-2 infection, hospitalization, and post-COVID-19 LTS provides valuable insights for public health interventions, resource allocation, and policy-making by identifying vulnerable populations, addressing disparities, and facilitating the implementation of targeted preventive measures.

Demographic and socio-economic risk factors of Hospital Admission, ICU admission and LTS.

Introduction

In the following section, we describe methods, results and conclusions to the analysis of socio-economics risk factors of LTS, hospital admission and ICU admission. We use the prospective cohort of hospitalised patients from WP2 to extend the analysis of clinical risk factors reported in WP2 D.2.6, by including as factors: ethnicity, education, labour situation, healthcare worker status, an indicator of laboratory-confirmed SARS-CoV-2 infection during the first wave, and cohort fixed effects.

Methods

Data Collection

The prospective cohort of SARS-CoV-2 infected patients from WP2 includes patients from 5 countries (Argentina, Italy, Spain, Netherlands, and France). The data collection was carried out using a custom-made structured electronic case report form (eCRF) developed in REDCap (Research Electronic Data Capture) and hosted at Consorzio Interuniversitario (CINECA). The variables underwent a process of homogenization across cohorts and standardization according to the WP2 protocol. The French COVID and COVID-HOME cohorts went through

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

a post-data collection harmonization process under the supervision of the Charité – Universitätsmedizin Berlin and data transformation conducted by the Centre Informatique National de l'Enseignement Superieur (CINES).

Five cohorts (French COVID, SAS, UBA, UNIBO and UNIVR) contribute to the analysis, including hospitalized and non-hospitalized patients aged >14 years old with a laboratory-confirmed SARS-CoV-2 infection. Patients were followed up at 3, 6, 12 and 18 months post-infection at an outpatient clinic setting. The enrollment took place at any of the above-mentioned time points, and the information collected included the socio-demographic characteristics of the patients.

In the WP2 ORCHESTRA cohort, baseline data includes symptom onset and positive SARS-CoV-2 test dates, demographic characteristics, comorbidities, clinical presentation, treatment during acute infection, and important outcomes like hospitalization, intensive care admission, and post-acute infection complications. These complications are categorized into respiratory, cardiac, embolic, neurological, acute renal, and gastrointestinal. Additionally, the twelve-month post-infection features cover symptoms, new medical events, vital signs, physical examination, laboratory parameters, and vaccination status.

Data Analysis

This analysis includes patients who completed the 12-month follow-up visit until 7th October 2022. This analysis is an extension of the WP2 D2.6, and thus, it applies the same LTS definition and considers the same set of clinical risk factors. Patients reporting at least one symptom at the 12-month follow-up assessment, either persisting since the acute infection or fluctuating/relapsing, were considered as having a LTS according to the WHO case definition [16]. In two additional analyses, we consider Hospital Admission and ICU admission as outcomes. The set of clinical risk factors considered in the LTS analysis are demographic (age and sex) and epidemiological data, hospitalization (hospitalized vs. outpatient management), ICU admission, clinical presentation during acute infection, treatment received including monoclonal antibodies, and vaccination status. In addition to the clinical factors which were part of WP2 D2.6, we extend the set of demographic factors to include ethnicity, education, labour situation and being a healthcare worker. In addition, we include an indicator of laboratory-confirmed SARS-CoV-2 infection during the first wave. As in WP2 D2.6, missing data from patients were not imputed, and therefore analysis was conducted using complete cases only.

Descriptive summary statistics between the outcome and the demographic variables are provided in Supplementary Material A. Means and standard deviations (SD) were calculated for continuous variables, and frequency tables and respective percentages were calculated for categorical variables. As part of the univariate analysis, crude odds ratios (OR) with 95% confidence intervals (95% CI) and their corresponding p-values are shown, where a p-value less than 0.05 was deemed as statically significant.

In the multivariate analysis, we run a logistic regression model, by considering a generalized linear model (GLM) with log-odds linking function, of LTS on demographic factors (conditional on the clinical factors from the WP2 D2.6 best model), and of Hospital Admission and ICU admission on demographic factors. All results include a specification in which cohort fixed effects were included in the analysis, which helps control for any unobserved differences across cohorts that could potentially confound the results.

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

In this report, the LTS outcome model was built after the WP2 D2.6 best model selection. In WP2 D2.6, the risk factors identified as statistically significant in the univariate analysis were selected to be included in the multivariate analysis. Model selection was done by evaluating the AIC (Akaike Information Criterion) of all possible combinations (subsets) of factors, and selecting the model with the lowest AIC score; adjusting for age and sex as additional risk factors.

Results

Descriptive analysis

Until 7th October 2022, 3668 patients completed the 12-month follow-up. In deliverable WP2 D2.6, the limit date was 30th June 2022, which explains the difference in the total number of patients that fit the inclusion criteria. Of the 3668 patients included in the analysis, 1796 had information on the LTS outcome and 3668 on the Hospital Admission and ICU Admission outcomes.

Overall, the majority of patients were male (2105; 57%) and aged between 40 and 60 years old (1441; 39%) and 60 and 80 years old (1509; 41%). The sample was predominantly Caucasian (2335; 80%), with most of the non-Caucasian population concentrated in the French COVID cohort (518; 91%). Patients with a high-school degree were the most frequent (846, 48%), followed by college graduates (522, 30%) and patients with no formal education (391, 22%). The majority of the sample reported to be employed or self-employed (1199; 66%), with 22% of patients unemployed (375) and 12% in informal employment (212). There were only 52 students (3%). Healthcare workers accounted for 338 individuals (11%).

In Supplementary material A, tables A, B and C show descriptive summary statistics between the demographic variables and the LTS, Hospital Admission and ICU admission outcomes, respectively. From the 1796 patients with no missing data for the LTS outcome, 1030 (57%) met the LTS definition after 12 months from acute infection. Female patients had a higher risk of persistence of symptoms after 12 months compared to men ($p<0.001$), and the Caucasian ethnic group had a lower presentation of LTS in our sample compared to all other ethnic groups ($p=0.005$). No differences were found according to age groups, education, labour situation and healthcare worker status.

In our sample, 2879 (78%) patients were hospitalised and 885 (24%) had an ICU admission. Table B shows that there were significant differences in the proportion of hospitalised patients across all demographic variables ($p<0.001$). There is an increasing trend in the probability of hospitalisation by age group, from 47% in those aged 15-30 to 94% in those aged >80. Hospitalisation was higher amongst male patients (85%) and non-Caucasians (97%) compared to females (70%) and Caucasians (67%), respectively. Concerning education, those with no formal education (77%) had a higher probability of hospitalisation than those with high school (48%) and college degrees (60%). Regarding the labour situation of individuals in our sample, the higher amount of hospitalised patients was amongst those informally employed (86%) followed by those employed and unemployed (58%). Students had the lowest number with only 6% of hospitalisations from a total of 52 students in the sample. Last, healthcare workers were less likely to be hospitalised than non-healthcare workers (72% vs. 81%). In Table C, from the Supplementary material, we observe a similar distribution of patients across categories for the ICU Admission outcome. There is an increasing trend in the probability of ICU Admission by age group, except for those >80 years old who have similar numbers to those aged 15-30 (6% and 9%, respectively). ICU Admission was higher amongst male

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

patients (31%), non-Caucasians (28%) and non-healthcare workers (25%) compared to females (14%), Caucasians (17%) and healthcare workers (16%), respectively. With regards to education, those with no formal education (77%) had a higher probability of ICU Admission, followed by those with a college (13%) and with high-school (12%) degree. Last, with respect to the labour situation, the higher amount of ICU-admitted patients was amongst those informally employed (22%) followed by those employed and unemployed (14% and 13%, respectively). There was only one ICU admission out of 52 students in the sample (2%).

Multivariable Analysis

In Tables 1, 2 and 3 we show the results of the multivariate analysis of LTS, hospital admission, and ICU admission, respectively, on socio-economic risk factors. The LTS results were conditioned on a set of relevant clinical factors. All results include a specification in which cohort fixed effects were included in the analysis.

LTS

Table 1 shows that no demographic or socio-economic factors, beyond female sex, are associated with LTS, after controlling for clinical risk factors and cohort fixed effects.

Variable	OR	95% CI		p-value
		Lower	Upper	
Female	1.705	1.207	2.409	**0.002**
Age	1.000	0.986	1.014	0.977
Caucasian	1.650	0.879	3.097	0.119
Education: High-school	0.731	0.459	1.165	0.188
Education: College	0.781	0.482	1.267	0.317
Labour: Informal employment	0.899	0.520	1.555	0.703
Labour: Student	0.467	0.161	1.352	0.160
Labour: Unemployed	1.266	0.789	2.032	0.329
Healthcare worker	0.683	0.409	1.138	0.143
Conditional on Clinical Factors		Yes		
Cohort Fixed Effects		Yes		
Number of observations	707			
Accuracy	0.650			
Balanced Accuracy	0.650			

Table 1. Multivariable analysis of demographic and socio-economic variables associated with the presence of at least one symptom at 12-month follow-up assessment. Accuracy: 65%.

Hospital admission

The results of the analyses displayed in Table 2 revealed several factors associated with the likelihood of hospital admission for COVID-19. Male sex and older patients, those who were infected during the first wave of the pandemic, and individuals in informal employment were more likely to be admitted to the hospital. On the other hand, results show that certain groups were less likely to be admitted to the hospital. These groups include students and healthcare workers. Results are consistent across specifications. We observe that the increased likelihood of hospitalisation for non-Caucasians is driven by cohort effects. This can be explained by the high concentration of non-caucasian individuals within the cohort of hospitalised patients from France (French COVID).

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

Variable	OR	95% CI Lower	95% CI Upper	p-value	OR	95% CI Lower	95% CI Upper	p-value
Female	0.447	0.337	0.592	**<0.001**	0.471	0.337	0.660	**<0.001**
Age	1.033	1.021	1.046	**<0.001**	1.030	1.013	1.048	**<0.001**
First Wave Infection	112.206	46.475	270.901	**<0.001**	27.221	9.544	77.635	**<0.001**
Fragile Population	1.129	0.835	1.527	0.430	1.164	0.804	1.685	0.421
Ethnic group: Caucasian	0.103	0.049	0.216	**<0.001**	0.617	0.213	1.789	0.374
Education: High-school	1.066	0.727	1.563	0.745	1.137	0.725	1.782	0.576
Education: College	0.927	0.600	1.433	0.733	0.707	0.407	1.230	0.220
Labour: Informal employment	2.404	1.404	4.116	**0.001**	2.040	1.051	3.961	**0.035**
Labour: Student	0.255	0.056	1.155	**0.076**	0.127	0.017	0.936	**0.043**
Labour: Unemployed	0.896	0.629	1.275	0.541	1.102	0.713	1.703	0.661
Healthcare worker	0.421	0.250	0.707	**0.001**	0.418	0.208	0.837	**0.014**
Conditional on Clinical Factors		No				No		
Cohort Fixed Effects		No				Yes		
Number of observations	1269				1269			
Accuracy	0.720				0.840			
Balanced Accuracy	0.740				0.860			

Table 2. Multivariable analysis of demographic and socio-economic variables associated with hospital admission. Accuracy: 86%.

ICU Admission

The results for the ICU admission outcome, shown in Table 3, are not consistent across specifications. After the inclusion of cohort fixed effects, only female sex is significantly associated with a decrease in the likelihood of ICU admission ($p<0.001$). Fragile populations (positive) and healthcare workers (negative) are weakly associated with ICU admission both before and after the inclusion of cohort fixed effects.

Variable	OR	95% CI Lower	95% CI Upper	p-value	OR	95% CI Lower	95% CI Upper	p-value
Female	0.372	0.255	0.544	**<0.001**	0.387	0.264	0.569	**<0.001**
Age	0.999	0.986	1.013	0.919	0.996	0.983	1.009	0.535
First Wave Infection	2.374	1.605	3.510	**<0.001**	1.193	0.737	1.930	0.473
Fragile Population	1.430	0.977	2.095	**0.066**	1.403	0.954	2.064	**0.085**
Ethnic group: Caucasian	0.547	0.328	0.912	**0.021**	0.741	0.440	1.248	0.259
Education: High-school	1.408	0.890	2.227	0.144	1.611	1.001	2.593	**0.050**
Education: College	1.157	0.715	1.871	0.553	1.115	0.686	1.813	0.661
Labour: Informal employment	1.627	1.006	2.630	**0.047**	1.462	0.894	2.391	0.130
Labour: Student	0.420	0.053	3.345	0.412	0.474	0.059	3.796	0.482
Labour: Unemployed	0.959	0.596	1.545	0.865	1.123	0.687	1.834	0.644
Healthcare worker	0.559	0.297	1.049	**0.070**	0.567	0.300	1.071	**0.080**
Conditional on Clinical Factors		No				No		
Cohort Fixed Effects		No				Yes		
Number of observations	1269				1269			
Accuracy	0.590				0.640			
Balanced Accuracy	0.600				0.630			

Table 3. Multivariable analysis of demographic and socio-economic variables associated with ICU admission. Accuracy: 63%.

Conclusions

- Female sex is associated with the persistence of symptoms at 12-month follow-up and a decreased likelihood of hospital and ICU admission.

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

- Healthcare workers were less likely of hospital admission, and, weakly, of ICU admission.
- Infections during the first half of the pandemic were the most likely to lead to hospitalisation.
- Some initial results were driven by cohort characteristics, for example, in the association between hospital admission and ethnicity. That is, the proportion of hospitalisations/ICU admissions and socio-economic characteristics of cohorts were the underlying drivers of the first specification results.
- The results of this analysis are subject to selection bias, as the probability of infection could not be accounted for in the analysis due to the use of a cohort of SARS-CoV-2 infected individuals.

Environmental risk factors of SARS-CoV-2 Infection, Antibody Response, COVID-19 Disease, and COVID-19 Vaccine Antibody Response in a General Population Cohort (COVICAT Study, Catalonia)

Introduction

We report on the work published by Kogevinas et al (2021) and Kogevinas et al (2023) about *"Ambient Air Pollution in Relation to SARS-CoV-2 Infection, Antibody Response, and COVID-19 Disease"* and *"Long-Term Exposure to Air Pollution and COVID-19 Vaccine Antibody Response"*, respectively, using a Cohort Study in Catalonia, Spain (COVICAT Study). This study was conducted prior to the COVICAT cohort joining ORCHESTRA-WP3 and the study was not funded by ORCHESTRA.

The first study, titled *"Ambient Air Pollution in Relation to SARS-CoV-2 Infection, Antibody Response, and COVID-19 Disease",* examines the potential association between ambient air pollution and SARS-CoV-2 infection, antibody response, and COVID-19 disease severity.

The article titled *"Long-Term Exposure to Air Pollution and COVID-19 Vaccine Antibody Response"* investigates the association between of long-term air pollution exposure and the antibody response to COVID-19 vaccines.

We describe data collection, variables, analytical approach, results and conclusions for both articles.

Methods

Data Collection

The COVICAT cohort (COVID-19 cohort in Catalonia) involves five existing adult cohort studies conducted prior to the COVID-19 pandemic. The largest cohort, Genomes for Life (GCAT), enrolled middle-aged individuals (40-65 years old) who are residents of Catalonia. Recruitment for GCAT began in 2015, primarily targeting blood donors through the Blood and Tissue Bank, a public agency. The study also includes four additional smaller cohorts: Multi Case-Control (MCC)-Spain, European Community Respiratory Health Survey (ECRHS), Urban Training, and Acute Kidney Injury in Agricultural Workers in Spain: Risk Factors and Long-Term Effects (LeRAgs).

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

Participants from the various cohorts were contacted and invited to participate to the COVICAT study. The majority of participants (n=8923) were from the GCAT cohort, while smaller numbers were included from the other cohorts (MCC-Spain: n=325, ECRHS: n=112, Urban Training and LeRAgs: n=55). The COVICAT study aimed to increase the proportion of older participants and those residing in rural areas. Harmonization of data from all cohorts was carried out after the outbreak of the COVID-19 pandemic in Spain in March 2020.

Data collection for COVICAT involved contacting participants via email or telephone and inviting them to respond to a questionnaire and provide a blood sample to determine SARS-CoV-2 seroprevalence. The questionnaire and blood sample collection took place between May 2020 and November 2020. The overall response rate among eligible participants was 61.6% (n=10 862). Air pollution exposure was evaluated for both responders and non-responders with valid pre-pandemic addresses.

For the vaccine response study, which was limited to the GCAT cohort, participants were re-contacted in spring 2021 after COVID-19 vaccination began in Spain. Questionnaires and blood samples were collected from a subset of participants (N=1090), out of 2404 who provided samples in 2020. A total of 927 vaccinated participants were included in this analysis after excluding individuals who were not vaccinated, received the Janssen vaccine, were below 40 years old, or had incomplete information.

For the GCAT cohort, a pre-pandemic follow-up was conducted in 2018-2019, and over 90% of participants who participated in that follow-up also participated in the COVICAT study. Data collection was primarily done through a study website where participants completed the questionnaire. Telephone administration of the questionnaire was provided for participants who preferred not to participate online (5.5% of the sample). All participants contacted from the cohort studies had consented in the past to be recontacted. Ethical approval for COVICAT was obtained from the Parc de Salut Mar Ethics Committee (CEIM-PS MAR, no. 2020/9307/I) and Hospital Universitari Germans Trias i Pujol Ethics Committee (CEI no. PI-20-182). All participants provided informed consent.

Variables

<u>Outcomes</u>

In the first study, the three primary outcomes were (i) quantitative serologically confirmed SARS-CoV-2 infection; (ii) antibody response based on IgM, IgA, and IgG levels, among those infected; and (iii) COVID-19 disease.

In the second study, the primary outcome was based on levels of IgG to spike proteins in non-infected vaccinated persons.

SARS-CoV-2 serology

All participants were invited to participate in a serological study, and 8906 participants agreed. From those who agreed, blood samples were collected from 4103 randomly selected participants. The blood samples were processed and analyzed to assess levels of IgM, IgA, and IgG antibodies using high-throughput multiplex quantitative suspension array technology.

In the second study a second follow-up of blood samples was analyzed for anti-SARS-CoV-2 antibody levels. IgG, IgM, and IgA levels were assessed using high-throughput multiplex

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

quantitative suspension array technology. The overall serostatus was defined by isotype and antigen. IgM results were specifically presented for participants who received one dose of the vaccine within one month of the 2021 study visit.

COVID-19 disease

COVID-19 disease definition included individuals who reported COVID-19 hospital admission, those with a prior positive diagnostic test for SARS-CoV-2 infection, and those with a combination of four or more COVID-19-related symptoms along with being in contact with a diagnosed COVID-19 case (n=481). This case definition was correlated with the presence of antibodies. SARS-CoV-2 antibodies were detected in a significant proportion of self-reported cases (70%) and participants with prior hospital admission (90%). Severe COVID-19 disease was defined as hospital admission, intensive care unit (ICU) admission, or receiving oxygen therapy without hospitalization.

Vaccination and SARS-CoV-2 Infection

The outcome analysis of the second study involved retrieving data from electronic health records regarding the number of vaccine doses, administration dates, and trade names of vaccines received by each participant. The vaccines administered were Comirnaty (BNT162b2, BioNTech-Pfizer), Spikevax (mRNA-1273, Moderna), and Vaxzevria (ChAdOx1 nCoV-19, Oxford–AstraZeneca). To identify participants who were previously infected with SARS-CoV-2, a two-part strategy was employed: (i) positive viral detection test results or self-reporting of positive tests, and (ii) seropositivity based on antibody data, including seropositivity in the pre-vaccination 2020 sample, or seropositivity to N-antigen in the 2021 sample.

Air Pollution Exposure

Air pollution exposure variables were generated by linking participants' addresses before the COVID-19 pandemic to estimates of long-term exposure to air pollutants, including particulate matter with an aerodynamic diameter of ≤2.5μm (PM2.5), black carbon (BC), nitrogen dioxide (NO2), and ozone (O3). The air pollution estimates for the period 2018-2019 were based on models developed by the Effects of Low-Level Air Pollution: A Study in Europe (ELAPSE) project, which utilized a combination of air pollution monitoring data, satellite observations, dispersion model estimates, land use, and traffic variables as predictors. Participants were assigned the 2010 annual average concentrations based on predicted surfaces at a resolution of 100x100 meters from the ELAPSE model. To estimate exposures for the years 2018 and 2019, a temporal correction was applied following protocols for temporal extrapolation developed in the ESCAPE project.

Covariates

The common list of covariates included in both analyses were demographic characteristics, such as age, sex, and educational level; prior diagnosis of chronic diseases; changes in residential address from the pre-pandemic period; and census-track data, including the deprivation index, population density, and degree of urbanization. Both studies also included lifestyle factors, like alcohol consumption and diet in the first study and physical activity in the second study. In addition, each study included a separate set of covariates. In the first study, symptoms related to COVID-19, such as fever, cough, and loss of odour/taste, were included in the analysis. In the second study, the list of covariates got extended by the inclusion of work-

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

related information during the pandemic, mask usage, mental health symptoms experienced during the pandemic, and exposure to green spaces.

Data Analysis

The first study applied log-binomial regression models to estimate risk ratios (RRs) and 95% confidence intervals (CIs) separately for infection and COVID-19 disease. We used linear regression models to estimate the association between air pollution and log10-transformed antibody levels among SARS-CoV-2 infected participants. Multinomial regression models were applied to allow for nominal dependent variables with more than two categories to estimate risk ratios (RRRs) and 95% CIs between air pollution and COVID-19 severity.

In the second study, the association between air pollution levels and log10 transformed antibody levels was estimated using linear regression models. The results are expressed as the percentage change in the geometric mean of the antibody levels along with 95% confidence intervals (95% CI). Air pollutants were modelled continuously and estimates per interquartile range (IQR) of each pollutant were reported. Stratified analyses were conducted based on participants' infection status, and adjustments were made for the severity of infection. Participants with missing covariates are excluded from the complete-case analysis models. Generalized additive models are used to explore the relationship between days since vaccination and IgG antibody levels among participants without prior infection, considering different levels of air pollution.

The statistical analyses were performed using Stata/SE (version 16; StataCorp LLC).

Results

The first study findings revealed that higher levels of ambient air pollution, specifically NO2 and PM2.5, were associated with an increased risk of SARS-CoV-2 infection. Additionally, individuals exposed to higher levels of air pollution exhibited lower antibody response to the virus, indicating potential implications for immune system function. Furthermore, the study found that elevated levels of air pollution, particularly PM2.5, were associated with an increased risk of developing severe COVID-19 disease. The association between air pollution and COVID-19 severity was stronger among older individuals and those with pre-existing health conditions, highlighting potential vulnerability factors.

In the second study, results suggest that long-term exposure to air pollution may have a detrimental effect on the antibody response to COVID-19 vaccines amongst non-infected individuals. The analysis revealed associations between certain air pollutants and reduced antibody response, even after adjusting for various confounding factors. Among vaccinated persons not infected by SARS-CoV-2 (n=632), higher pre-pandemic air pollution levels were associated with lower vaccine antibody response for IgM (one-month post-vaccination) and IgG. Percentage change in geometric mean IgG levels per interquartile range of PM2.5 (1.7μg/m3) were -8.1 (95%CI -15.9, 0.4) for RBD, -9.9 (-16.2, -3.1) for S and -8.4 (-13.5, -3.0) for S2. A similar pattern was observed for NO2 and BC and an inverse pattern for O3. Differences in IgG levels by air pollution levels persisted with time since vaccination. The results did not show an association of air pollution with vaccine antibody response among participants with prior infection (n=295).

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

Conclusion

The COVICAT Study provides valuable insights into the relationship between air pollution and SARS-CoV-2 infection, antibody response, COVID-19 disease severity, and COVID-19 Vaccine Antibody Response.

The findings from the article *"Ambient Air Pollution in Relation to SARS-CoV-2 Infection, Antibody Response, and COVID-19 Disease"* suggest that reducing air pollution levels may reduce the risk of SARS-CoV-2 infection and its severity.

The study *"Long-Term Exposure to Air Pollution and COVID-19 Vaccine Antibody Response"* shows a potential link between air pollution and reduced immune response to COVID-19 vaccination.

Overall, these findings emphasize the importance of considering air pollution as a potential factor in evaluating vaccine effectiveness and public health outcomes.

References

[1] Platt L, and Warwick R. (2020). COVID-19 and Ethnic Inequalities in England and Wales*. Fiscal Studies. 41(2): p. 259-289.

[2] Blundell R, et al. (2020). *COVID-19 and Inequalities**. Fiscal Studies. 41(2): p. 291-319.

[3] Patel JA, et al. (2020). *Poverty, inequality and COVID-19: the forgotten vulnerable.* Public Health. 183: p. 110-111.

[4] Plumper T, and Neumayer E. (2020). *The pandemic predominantly hits poor neighbourhoods? SARS-CoV-2 infections and COVID-19 fatalities in German districts.* European Journal of Public Health. 30(6): p. 1176-1180.

[5] Bowe B, Xie Y, Gibson AK, Cai M, van Donkelaar A, Martin RV, et al. (2021). Ambient fine particulate matter air pollution and the risk of hospitalization among COVID-19 positive individuals: Cohort study. Environment International 154:106564; doi:10.1016/j.envint.2021.106564.

[6] Chen Z, Sidell MA, Huang BZ, Chow T, Eckel SP, Martinez MP, et al. (2022a). Ambient Air Pollutant Exposures and COVID-19 Severity and Mortality in a Cohort of COVID-19 Patients in Southern California. Am J Respir Crit Care Med; doi:10.1164/rccm.202108-1909OC.

[7] Yu Z, Bellander T, Bergström A, Dillner J, Eneroth K, Engardt M, et al. (2022). Association of Short-term Air Pollution Exposure With SARS-CoV-2 Infection Among Young Adults in Sweden. JAMA Netw Open 5:e228109; doi:10.1001/jamanetworkopen.2022.8109.

[8] Nobile F, Michelozzi P, Ancona C, et al. (2022). Air pollution, SARS-CoV-2 incidence and COVID-19 mortality in Rome – a longitudinal study. Eur Respir J. 2200589. doi:10.1183/13993003.00589-2022

[9] English PB, Von Behren J, Balmes JR, et al. (2022). Association between long-term exposure to particulate air pollution with SARS-CoV-2 infections and COVID-19 deaths in California, U.S.A. Environmental Advances. 9:100270. doi:10.1016/j.envadv.2022.100270

[10] Kim H, Samet JM, Bell ML. (2022). Association between Short-Term Exposure to Air Pollution and COVID-19 Mortality: A Population-Based Case-Crossover Study Using Individual-Level Mortality Registry Confirmed by Medical Examiners. Environ Health Perspect. 130(11):117006. doi: 10.1289/EHP10836. PMID: 3636778

[11] Chen C, Wang J, Kwong J, Kim J, van Donkelaar A, Martin RV, et al. (2022b). Association between long-term exposure to ambient air pollution and COVID-19 severity: a prospective cohort study. CMAJ 194:E693–E700; doi:10.1503/cmaj.220068.

[12] Sheridan C, Klompmaker J, Cummins S, et al. (2022). Associations of air pollution with COVID-19 positivity, hospitalisations, and mortality: Observational evidence from UK Biobank. Environmental Pollution. 308:119686. doi:10.1016/j.envpol.2022.119686

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

[13] Villeneuve PJ, Goldberg MS. (2020). Methodological Considerations for Epidemiological Studies of Air Pollution and the SARS and COVID-19 Coronavirus Outbreaks. Environ Health Perspect 128:095001; doi:10.1289/EHP7411.

[14] Kogevinas M, Castaño-Vinyals G, Karachaliou M, Espinosa A, de Cid R, Garcia-Aymerich J, et al. (2021). Ambient Air Pollution in Relation to SARS-CoV-2 Infection, Antibody Response, and COVID-19 Disease: A Cohort Study in Catalonia, Spain (COVICAT Study). Environ Health Perspect 129:117003; doi:10.1289/EHP9726.

[15] Kogevinas, M., Karachaliou, M., Espinosa, A., Aguilar, R., Castaño-Vinyals, G., Garcia-Aymerich, J., Carreras, A., Cortés, B., Pleguezuelos, V., Papantoniou, K., Rubio, R., Jiménez, A., Vidal, M., Serra, P., Parras, D., Santamaría, P., Izquierdo, L., Cirach, M., Nieuwenhuijsen, M., Dadvand, P., … Tonne, C. (2023). Long-Term Exposure to Air Pollution and COVID-19 Vaccine Antibody Response in a General Population Cohort (COVICAT Study, Catalonia). *Environmental health perspectives*, *131*(4), 47001. https://doi.org/10.1289/EHP11989

[16] WHO. A clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021. Available online: https://www.who.int/publications/i/item/WHO-2019-nCoV-Post_COVID-19_condition-Clinical_case_definition-2021.1 (accessed on 24 June 2022).

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167

Supplementary material

A. Tables

Table A. Summary Statistics of post-COVID-19 LTS by Demographic Factors: Number of non-missing information (N), Frequency, Proportion (%), Odds-ratio (OR), lower and upper 95% CI bound and p-value.

Variable	N	Frequency	Proportion	Crude OR	95% LB	95% UB	p-value
Age (N=1,794)							
15-30	98	50	51.02%	1.00			
31-40	150	83	55.33%	1.19	0.71	1.99	0.509
41-60	774	456	58.91%	1.38	0.90	2.10	0.139
61-80	689	403	58.49%	1.35	0.88	2.07	0.165
>80	83	38	45.78%	0.81	0.45	1.46	0.487
Sex (N=1,795)							
Male	1016	536	52.76%	1.00			
Female	779	493	63.29%	1.54	1.28	1.87	**<0.001**
Ethnicity = Caucasian (N=1,531)							
No	261	166	63.60%	1.00			
Yes	1270	688	54.17%	0.68	0.51	0.89	**0.005**
Education (N=1,058)							
No formal education	198	109	55.05%	1.00			
High-school	561	273	48.66%	0.77	0.56	1.07	0.123
College	299	161	53.85%	0.95	0.66	1.37	0.793
Labour Situation (N=1,064)							
Employed	693	374	53.97%	1.00			
Informally employed	110	53	48.18%	0.79	0.53	1.19	0.261
Student	38	16	42.11%	0.62	0.31	1.20	0.159
Unemployed	223	114	51.12%	0.89	0.66	1.21	0.460
Healthcare Worker (N=1,509)							
No	1347	797	59.17%	1.00			
Yes	162	93	57.41%	0.93	0.67	1.30	0.666

Table B. Summary Statistics of Hospital Admission by Demographic Factors: Number of non-missing information (N), Frequency, Proportion (%), Odds-ratio (OR), lower and upper 95% CI bound and p-value.

Variable	N	Frequency	Proportion	Crude OR	95% LB	95% UB	p-value
Age (N=3,666)							
15-30	187	88	47.06%	1.00			
31-40	298	209	70.13%	2.63	1.81	3.86	**<0.001**
41-60	1441	1067	74.05%	3.21	2.35	4.38	**<0.001**
61-80	1509	1297	85.95%	6.87	4.98	9.49	**<0.001**
>80	231	218	94.37%	18.54	10.20	36.38	**<0.001**
Sex (N=3,667)							
Male	2105	1789	84.99%	1.00			
Female	1562	1090	69.78%	0.41	0.35	0.48	**<0.001**

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101016167

Ethnicity = Caucasian (N=2,902)							
No	567	548	96.65%	1.00			
Yes	2335	1573	67.37%	0.07	0.04	0.11	**<0.001**
Education (N=1,760)							
No formal education	392	300	76.53%	1.00			
High-school	846	402	47.52%	0.28	0.21	0.36	**<0.001**
College	522	312	59.77%	0.46	0.34	0.61	**<0.001**
Labour Situation (N=1,838)							
Employed	1199	700	58.38%	1.00			
Informally employed	212	182	85.85%	4.30	2.92	6.56	**<0.001**
Student	52	3	5.77%	0.05	0.01	0.13	**<0.001**
Unemployed	375	219	58.40%	1.00	0.79	1.27	0.996
Healthcare Worker (N=3,183)							
No	2845	2315	81.37%	1.00			
Yes	338	243	71.89%	0.59	0.45	0.76	**<0.001**

Table C. Summary Statistics of ICU Admission by Demographic Factors: Number of non-missing information (N), Frequency, Proportion (%), Odds-ratio (OR), lower and upper 95% CI bound and p-value.

Variable	N	Frequency	Proportion	Crude OR	95% LB	95% UB	p-value
Age (N=3,666)							
15-30	187	12	6.42%	1.00			
31-40	298	49	16.44%	2.84	1.51	5.76	**<0.001**
41-60	1441	342	23.73%	4.48	2.57	8.61	**<0.001**
61-80	1509	462	30.62%	6.35	3.65	12.19	**<0.001**
>80	231	20	8.66%	1.37	0.66	2.99	0.401
Sex (N=3,667)							
Male	2105	660	31.35%	1.00			
Female	1562	225	14.40%	0.37	0.31	0.44	**<0.001**
Ethnicity = Caucasian (N=2,902)							
No	567	158	27.87%	1.00			
Yes	2335	404	17.30%	0.54	0.44	0.67	**<0.001**
Education (N=1,760)							
No formal education	392	65	16.58%	1.00			
High-school	846	99	11.70%	0.67	0.48	0.94	**0.021**
College	522	66	12.64%	0.73	0.50	1.06	0.095
Labour Situation (N=1,838)							
Employed	1199	171	14.26%	1.00			
Informally employed	212	46	21.70%	1.67	1.15	2.39	**0.008**
Student	52	1	1.92%	0.13	0.01	0.61	**0.004**
Unemployed	375	49	13.07%	0.91	0.64	1.26	0.566
Healthcare Worker (N=3,183)							
No	2845	698	24.53%	1.00			
Yes	338	55	16.27%	0.60	0.44	0.80	**<0.001**

ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101016167