

RESEARCH

Open Access



Risk, determinants, and persistence of long-COVID in a population-based cohort study in Catalonia

Manolis Kogevinas^{1,2,3,4*†}, Marianna Karachaliou^{1†}, Ana Espinosa^{1,2,3}, Susana Iraola-Guzmán^{5,6}, Gemma Castaño-Vinyals^{1,2,3}, Laura Delgado-Ortiz^{1,2,3}, Xavier Farré^{5,6}, Natàlia Blay^{5,6}, Neil Pearce⁷, Magda Bosch de Basea¹, Eva Alonso Nogués⁸, Carlota Dobaño^{1,9}, Gemma Moncunill^{1,9}, Rafael de Cid^{5,6} and Judith Garcia-Aymerich^{1,2,3}

Abstract

Background Long-COVID has mostly been investigated in clinical settings. We aimed to assess the risk, subtypes, persistence, and determinants of long-COVID in a prospective population-based study of adults with a history of SARS-CoV-2 infection in Catalonia.

Methods We examined 2764 infected individuals from a population-based cohort (COVICAT) established before the pandemic and followed up three times across the pandemic (2020, 2021, 2023). We assessed immunoglobulin (Ig)G levels against SARS-CoV-2, clinical, vaccination, sociodemographic, and lifestyle factors. Long-COVID risk and subtypes were defined based on participant-reported symptoms and electronic health records. We identified a total of 647 long-COVID cases and compared them with 2117 infected individuals without the condition.

Results Between 2021 and 2023, 23% of infected subjects developed long-COVID symptoms. In 56% of long-COVID cases in 2021, symptoms persisted for 2 years. Long-COVID presented clinically in three subtypes, mild neuromuscular, mild respiratory, and severe multi-organ. The latter was associated with persistent long-COVID. Risk was higher among females, participants under 50 years, of low socioeconomic status, severe COVID-19 infection, elevated pre-vaccination IgG levels, obesity, and prior chronic disease, particularly asthma/chronic obstructive pulmonary disease and mental health conditions. A lower risk was associated to pre-infection vaccination, infection after omicron became the dominant variant, higher physical activity levels, and sleeping 6–8 h. Vaccination during the 3 months post-infection was also protective against long-COVID.

Conclusions Long-COVID persisted for up to 2 years in half of the cases, and risk was influenced by multiple factors.

Keywords Long-COVID, Epidemiology, COVID-19, SARS-CoV-2 antibodies, Vaccines, Exercise, Sleep

[†]Manolis Kogevinas and Marianna Karachaliou share equal contribution.

*Correspondence:

Manolis Kogevinas
manolis.kogevinas@isglobal.org

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Long-term sequelae following acute infectious illnesses are well documented [1]. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been linked to a chronic disease state following the acute infection [2]. This condition, known as long-COVID, encompasses symptoms as diverse as fatigue, shortness of breath, cognitive dysfunction, digestive symptoms, depression, and sleep disturbances, among others. In 2021, the World Health Organization (WHO) defined long-COVID as symptoms persisting for at least two months, not attributable to an alternative diagnosis, 3 months post diagnosis [3]. Recently, the National Academies of Sciences, Engineering, and Medicine (NASEM) recognized long-COVID as an infection-associated chronic condition present for at least 3 months as a continuous, relapsing and remitting, or progressive disease state that affects one or more organ systems [4].

Estimates of long-COVID frequency vary due to the heterogeneous definitions and the predominance of clinical-based rather than population-based studies [1]. Among clinical-based samples, a pooled analysis of 54 studies (1.2 million individuals) estimated that 6.2% experience long-COVID symptoms [2]. The US RECOVER study, covering population-based and convenience sampling, estimated a 10.0% risk of long-COVID 6 months post-infection [3]. A population-based Dutch study found 12.7% of COVID-19 patients had persistent somatic symptoms [5] and in 10 UK longitudinal studies the prevalence of ongoing symptoms for +12 weeks was 7.8–17% [6]. The risk of long-COVID might change due to the circulation of different variants [7] and repeated infections [8, 9].

The mechanisms leading to long-COVID are multiple and still unclear [10–12]. Identifying subtypes and disease courses over time may help disentangle these mechanisms.

Research has consistently found that middle age, female sex, presence of chronic conditions, severity of initial COVID-19 infection and hospitalization, number of SARS-CoV-2 infections, and lack of vaccination relate to higher risk of long-COVID [13, 14], but scarce data exist on the role of these factors on long-COVID progression. Moreover, data is lacking about the potential role of other clinical, environmental, and lifestyle factors. Evidence comes mostly from clinical rather than population-based studies without sufficient follow-up time to allow identifying factors related to the long-term course of the disease.

We examined a population-based cohort with multiple clinical, sociodemographic, and lifestyle characteristics and SARS-CoV-2 multiplex serology available in three consecutive follow-ups (2020, 2021, and 2023),

along with electronic health records. We assessed the risk, course, subtypes, and risk factors of long-COVID, by comparing SARS-CoV-2 infected individuals experiencing long-COVID symptoms, with those without such condition.

Methods

Design and study population

This study is based on an adult COVID-19 population-based cohort network in Catalonia (COVICAT) [15–18] and focuses on participants recruited from the GCAT|Genomes for life cohort (www.genomesforlife.com) [19]. The GCAT cohort was established before the pandemic with a focus on investigating the etiology of chronic diseases. Participants from GCAT were invited to complete COVID-19-related surveys and provide samples in June–November 2020 (baseline, $n=8923$), in June–August 2021 (first follow-up, $n=7015$), and in February–May 2023 (second follow-up, $n=5215$, participation rate 74%). In this analysis, we included participants with SARS-CoV-2 infection at anytime during the study period who completed all three surveys ($n=2764$) (Fig. 1). We invited all eligible GCAT participants to complete the COVID-19-related surveys. Invitation was sent through an email (initial and two reminders). Participants without registered email addresses were contacted through the telephone. All participants contacted had consented in the past to be re-contacted. Data collection was primarily completed on a study website where participants filled in the questionnaire. We administered the questionnaire via telephone when participants were not comfortable with online study participation. Participants of the baseline and first follow-up who did not participate in the second follow-up were slightly older, and had a lower proportion of individuals with university education and with excellent/good health condition than those with a second follow-up (Additional File 1: Table S1). Percentage of vaccinated was slightly higher among participants (92.6%) compared to non-participants (88.5%). Differences, however, were minor for COVID-19 severity (6.0% non-participants vs 6.4% participants) or in long-COVID subtypes (27.0% vs 27.8% for severe multi-organ type, see below for definitions).

Ethical approval was obtained by the ethics committees at the Hospital Universitari Germans Trias i Pujol (CEI no. PI-20–182) and the Parc de Salut Mar (CEIM-PS MAR, no. 2020/9307/I). All participants provided informed consent and agreed to potential re-contact.

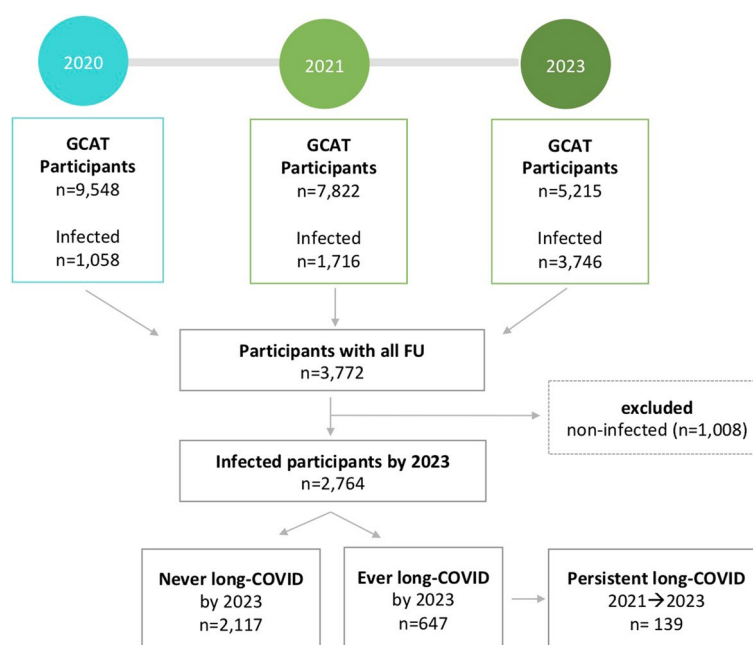


Fig. 1 Flowchart of the study population (COVICAT study), 2020 to 2023. See text for the definition of persistent long-COVID

Study procedures

In all study visits, participants answered validated questionnaires aligned with European harmonization and standardization guidelines [20], available at the study website, (<http://www.gcatbiobank.org/media/upload/arxius/COVICAT/encuesta%20COVICAT.pdf>), and provided blood samples. Participants' data was also linked to electronic health records of the Catalan universal coverage health system.

Blood samples and serology

Blood samples were collected in EDTA and PST tubes at the Banc de Sang i Teixits (Catalonia), stored at 4 °C, and processed in the GCAT lab (IGTP, Badalona) within 24 h after collection. From the EDTA blood sample, the levels of immunoglobulin G (IgG), M (IgM), and A (IgA) were assessed by high-throughput multiplex quantitative suspension array technology against a panel of five SARS-CoV-2 antigens: the spike full-length protein (S) and the receptor binding domain (RBD), the sub-region S2, the nucleocapsid full length (NFL) and the specific C-terminal (NCT) region [21, 22]. The serological assays were performed at the Immune Response and Biomarkers Core Facility at ISGlobal. We used serology to define SARS-CoV-2 infection (see below for details). We also used pre-vaccine serological data (baseline, 2020) to identify early biomarkers of humoral immunity induced by natural infection that contribute

to the later development of long-COVID (subsample of 1672 participants).

SARS-CoV-2 infection, COVID-19 severity, and long-COVID definitions

We defined SARS-CoV-2 infection as (i) self-reported: positive viral detection test (polymerase chain reaction or antigen test), COVID-19 disease, hospitalization or intensive care unit (ICU) admission due to COVID-19; (ii) identification of a positive test through record linkage with the repository of SARS-CoV-2 tests of the Emergency EHRs; (iii) identification of a diagnosis, hospitalization or ICU admission due to COVID-19 disease (ICD-10 codes U071, U072, U10, or J1282) through record linkage with EHRs; or (iv) SARS-CoV-2 seropositivity for the subsample with serological testing from any time (2020–2023). Seropositivity was based on our antibody data using the following criteria: seropositivity to any SARS-CoV-2 antigen in a pre-vaccination serology sample, or seropositivity to Nucleocapsid-antigen in a sample of vaccinated individuals, since the available vaccines did not contain or produce Nucleocapsid-antigen. In our study population ($n=2764$), 1492 (54%) fulfilled more than one criterion, 1171 (42.4%) were only detected through questionnaires, 79 (2.8%) were only detected by serology, and 22 (0.8%) were only detected by EHR. The date of SARS-CoV-2 infection accompanied the aforementioned diagnoses. There were 79 infected asymptomatic people identified only through serology for which

the date of infection was not possible to estimate but were allocated as being infected in specific time intervals using repeated serology in 2020, 2021, and 2023.

Using the date of first infection (earliest occurrence of infections as described above), we classified our population as first infected with pre-omicron or omicron variants. Thirty-five of the 79 individuals identified as infected based on serology could not be definitively classified as having been infected before or during the Omicron wave and were excluded from this analysis. We considered omicron infections all those occurring after November 29, 2021, as this corresponds to the date of the first omicron case in Catalonia according to the SARS-CoV-2 submitted sequences to GISAID (<https://www.gisaid.org>; up to 20 April 2023). We classified the severity of SARS-CoV-2 infection according to WHO as asymptomatic, mild/moderate, and severe/critical [23].

We defined long-COVID following the WHO and recent publications [24–27]: (i) the presence of SARS-CoV-2 infection (defined above) and (ii) a self-report of at least one of the following symptoms or sequels for a duration of at least three months. This definition aligns with the 2024 NASEM definition of long-COVID [4] as it considers a larger duration of symptoms (three instead of two months), compared to that of WHO. The specific questions on long-COVID symptoms in our 2021 and 2023 surveys (Additional File 1: Table S2) were harmonized and included: fever or low-grade fever, loss of appetite, dry cough, dyspnoea/shortness of breath, persistent chest pain/tightness, throat pain, other respiratory problems, tiredness/unusual fatigue, confusion/loss of speech/loss of movement, chills/vertigo/dizziness, loss of taste or smell, headache, other cognitive or neurological problems, conjunctivitis, other problems with dry eyes/mouth, diarrhea, nausea/vomiting, skin rashes, other dermatological problems, muscle/joint pain, other muscular problems, hematological problems, endocrine problems, psychological or psychiatric problems, cardiovascular problems, menstrual problems, renal problems. We also considered the identification of COVID-19 sequelae (ICD-10 B94.8) or long-COVID diagnosis (U09.9) from electronic health records during the study period. In 2023, we included a question on the presence of long-COVID symptoms during the last week, among those reporting long-term symptomatology. Non-infected participants were not screened for long-term symptoms.

We classified the long-COVID cases combining the reports of symptoms in the first (2021) and second (2023) follow-ups as: never long-COVID ($n=2117$, no long-COVID symptoms reported in 2021 or 2023); recovered long-COVID ($n=109$, long-COVID symptoms in 2021 without symptoms in 2023); incident long-COVID in 2023 ($n=399$, no long-COVID symptoms reported in

2021 and symptoms only reported in 2023); persistent long-COVID ($n=139$, long-COVID symptoms reported in 2021 and 2023). The 2023 interview covered a span of 2 years of follow-up. To ensure that symptoms in the persistent long-COVID group were present in 2023, we only included those reporting symptoms during the last week before the survey and classified those not having symptoms during the last week ($n=109$) as recovered. We considered the presence of any symptom recognizing that long-COVID presents as a continuous, relapsing and remitting, or progressive disease state.

Vaccination

We retrieved information on vaccinations from electronic health records. Participants received the following vaccines: Comirnaty (BNT162b2, mRNA, BioNTech-Pfizer, Mainz/New York City), Spikevax (mRNA-1273, Moderna, Cambridge), and Vaxzevria (ChAdOx1 nCoV-19, Oxford–AstraZeneca), and in smaller numbers the Janssen COVID-19 vaccine (Ad26.COV2.S, Johnson & Johnson–Janssen).

Other participant characteristics

We obtained information on individual characteristics: (i) socio-demographics (sex, age, education); (ii) deprivation index, linking pre-pandemic residential addresses to the 2011 census tract-level deprivation index [28], (iii) lifestyle (tobacco smoking, alcohol intake, physical activity levels according to the International Physical Activity Questionnaire, and sleeping duration); (iv) self-appraisal of health status, and (v) self-reported presence of chronic diseases (asthma, chronic bronchitis/chronic obstructive pulmonary disease (COPD)/emphysema, other respiratory diseases, hypertension, heart disease, diabetes, digestive diseases, liver diseases, gynecological diseases, joint problems/rheumatism, rheumatoid arthritis, other autoimmune diseases, HIV/other immunodeficiency problems, cancer, depression, anxiety, other mental health problems/addictions, kidney disease).

Statistical analysis

To identify subtypes of long-COVID reported in 2021 (first follow-up), we performed latent class analysis. We used all variables of symptoms at the same time, tested from one to seven latent classes, and selected the number of classes with the lowest Bayesian information criterion. Each participant was assigned to the latent cluster (here defined as long-COVID subtype) for which he/she had the highest membership probability.

To identify risk factors associated with long-COVID, we used Poisson regression models with robust variance, adjusting for age, sex, and education to compare SARS-CoV-2 infected individuals with and without

long-COVID. A multivariable stepwise model was constructed incorporating all identified predictors from prior analyses. We kept in the multivariable regression model the specific chronic diseases instead of the combined variable “any chronic disease” to help interpretation in relation to people with specific chronic diseases. These potential predictors included a comprehensive list of characteristics during the follow-up period preceding the first diagnosis of SARS-CoV-2. For individuals identified with COVID-19 in the 2021 follow-up, we utilized information from the 2020 questionnaire. Similarly, for those diagnosed in 2023, data from the 2021 questionnaire were employed. Due to the high vaccination rates in our population, we focused on serological determinants related to natural infection measured in 2020 (pre-vaccination) [15]. In an alternative analysis, we applied the false discovery rate correction [29], to identify significant results while limiting the proportion of false positives among them, ensuring a more reliable interpretation of findings. Predictors with q -values < 0.05 were considered significant under a false discovery rate of 5%.

Analyses were performed using Stata (version 16.1; Stata Corp LP; College Station, TX, USA), and the R

statistical program (version 4.3.0; R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was set at a p -value < 0.05 .

Results

Prevalence and clinical presentation of long-COVID

Among 2764 SARS-CoV-2 infected persons, 647 (23.4%) reported long-COVID symptoms (Fig. 1). The most common symptoms of long-COVID were neurological symptoms (63%), muscular problems (39%), respiratory symptoms (28%) and psychological and psychiatric symptoms (21%) (Additional File 1: Table S3). Neurological and muscular symptoms were more common in women while respiratory symptoms were more common in men. Long-COVID participants with the first infection during the omicron period presented with similar but fewer long-term symptoms than those with the first infection in the pre-omicron period (Additional File 1: Fig. S1).

We identified three subtypes of long-COVID in 2021 from latent class analysis (Fig. 2). The first ($n=128$, 51.6%) had a high prevalence of neurological symptoms and moderate prevalence of respiratory, skin,

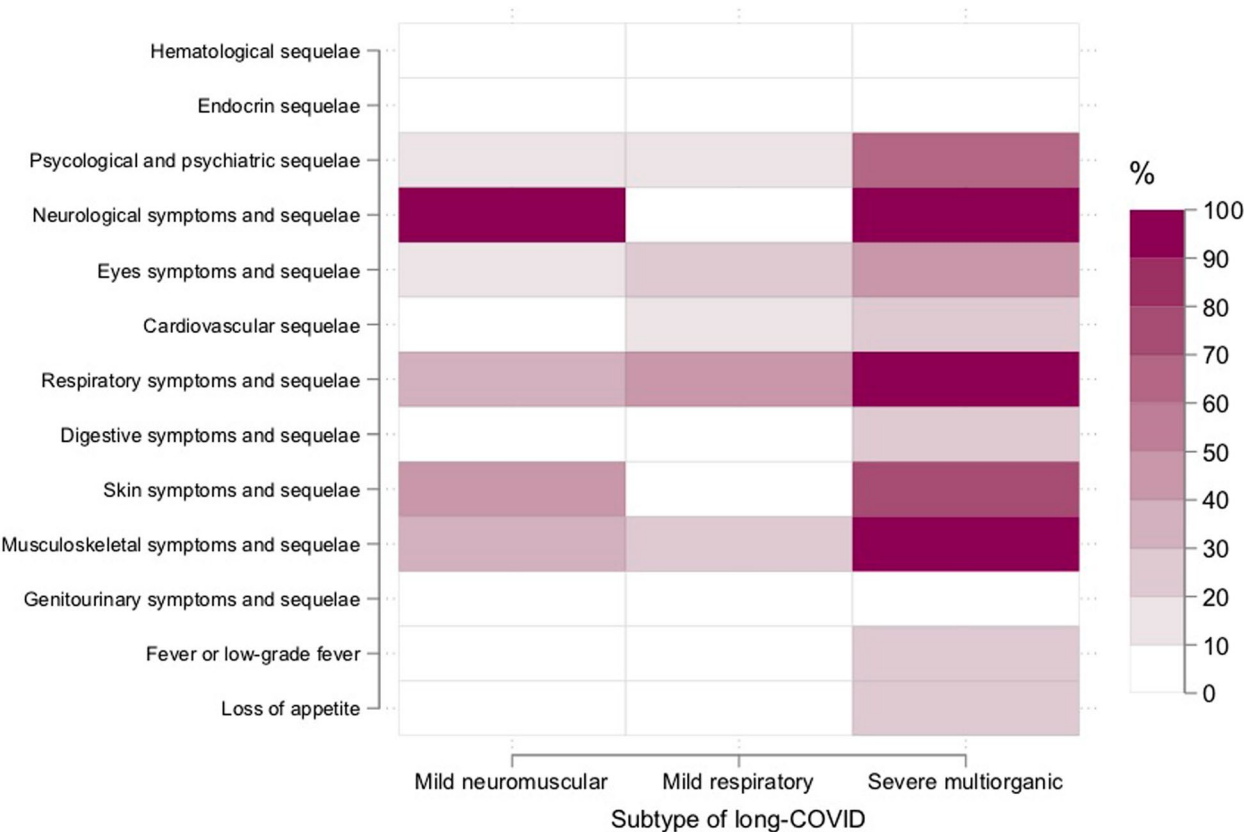


Fig. 2 Grouping of long-COVID symptoms in 2021 ($n=248$) by latent class analysis. Three subtypes are defined: mild neuro/muscular symptoms ($n=128$), mild respiratory symptoms ($n=51$), severe multi-organ symptoms ($n=69$)

musculoskeletal, and sensory symptoms, and was therefore labeled “mild neuromuscular.” The second ($n=51$, 20.6%) exhibited the highest prevalence of respiratory symptoms with a low-to-moderate prevalence of some additional symptoms and was labeled “mild respiratory.” The third ($n=69$, 27.8%) showed a high prevalence of psychological, neurological, respiratory, skin, and musculoskeletal symptoms and low-to-moderate prevalence of all remaining symptoms and was labeled “severe multi-organ.”

Among cases of long-COVID in 2021 ($n=248$), 56.0% reported persistent symptoms in 2023 (persistent long-COVID cases), while 44.0% indicated no recent symptoms in 2023 (recovered long-COVID). Among the long-COVID subtypes identified in 2021, the highest risk for persistent long-COVID was observed for the severe multi-organ subtype ($RR=1.61$, 1.13–2.28) (Table 1).

Sociodemographic factors

Long-COVID risk was higher in women ($RR=1.47$, 95% CI 1.26–1.71) than in men and in persons with primary or lower education ($RR=1.31$, 95% CI 1.03–1.67) compared to those with university education (Table 2). The risk was lower in persons above 65 years ($RR=0.77$, 95% CI 0.61–0.98) compared to those younger than 50 years of age.

Lifestyle factors and obesity

Among lifestyle factors (Table 2), we observed a protective association between physical activity levels and long-COVID both for moderate to high ($RR=0.80$, 95% CI 0.67–0.96, high vs. low physical activity). There was a U-shaped relationship with sleep duration. Those with ≤ 5 ($RR=1.31$, 95% CI 1.06–1.62) or ≥ 9 h sleep ($RR=1.74$, 95% CI 1.30–2.34) had higher risks than those

with 7 h duration. Obesity ($BMI \geq 30$ kg/m²) was associated with a higher risk of long-COVID ($RR=1.39$, 95% CI 1.20–1.62).

Prior medical history

Participants reporting regular/poor perceived health in 2020 had an increased risk of long-COVID ($RR=2.26$, 95% CI 1.89–2.71) compared to those reporting very good/excellent health (Table 3). Presence of any prior chronic condition was associated with a higher risk of long-COVID ($RR=1.58$, 95% CI 1.38–1.80). Among specific major chronic diseases, an increased risk was observed for nearly all diseases, with highest risks for prior anxiety/depression ($RR=1.95$, 95% CI 1.58–2.42), digestive diseases ($RR=1.79$, 95% CI 1.41–2.25), joint problems/rheumatisms ($RR=1.69$, 95% CI 1.39–2.06), and for chronic respiratory diseases (asthma or COPD) ($RR=1.45$, 95% CI 1.21–1.73).

COVID-19 severity, variants, and vaccines

The severity of infection was strongly associated with the risk of long-COVID (Table 4). Participants with mild/moderate infection had three times higher risk ($RR=3.10$, 95% CI 2.20–4.36), and those with severe/critical infection more than nine times higher risk ($RR=9.88$, 95% CI 6.88–14.18) of long-COVID, compared with those asymptomatically infected.

First infection during the omicron period was associated with a lower risk of long-COVID compared to earlier variants ($RR=0.36$, 95% CI 0.32–0.42). We additionally adjusted for vaccination prior to the first infection and estimates for omicron continue being clearly lower than those for pre-omicron although less extreme ($RR=0.56$, 95% CI 0.43–0.73). Although the regression models converged, there was collinearity between the two variables, pre/omicron and vaccination, with a high variance inflation factor (> 10). The determinants of long-COVID identified were generally consistent for the pre-omicron and omicron variants apart from the association of severity of infection that was weaker for omicron-related long-COVID, whereas health status, any chronic disease, and asthma/COPD associations were more pronounced for omicron related long-COVID (Additional File 1: Table S4). Smoking emerged as a risk factor for omicron-related long-COVID (Additional File 1: Table S4).

Strong protection for long-COVID was observed for vaccination prior to infection ($RR=0.33$, 95% CI 0.29–0.38) (Table 4). Repeated vaccination prior to infection was protective for long-COVID with the lowest risks observed for those with ≥ 4 doses. Vaccination within 3 months post-infection was also protective against long-COVID ($RR=0.58$, 95% CI 0.39–0.86).

Table 1 Long-COVID subtypes in 2021 and their risk of persistence up to 2023 COVICAT cohort ($n=248$)

	Recovered up to 2023 $n=109$ n (%)	Persistent up to 2023 $n=139$ n (%)	RR (95% CI) ^b
Long-COVID subtypes in 2021 ^a			
Mild neuromuscular	30 (27.5)	21 (15.1)	ref
Mild respiratory	61 (56.0)	67 (48.2)	1.22 (0.85,1.74)
Severe multi-organ	18 (16.5)	51 (36.7)	1.61 (1.13,2.28)

RR relative risk, CI confidence intervals

^a Defined by latent class analysis of symptoms and sequels reported in 2021 (see Fig. 2)

^b Adjusted for age, sex, and educational level

Table 2 Sociodemographic and lifestyle characteristics and their association (RR, (95% CI)) with long-COVID development and persistence in 2764 SARS-CoV-2-infected participants of the COVICAT cohort, 2020–2023

	Never long-COVID <i>n</i> = 2117 <i>n</i> (%)	Ever long-COVID <i>n</i> = 647 <i>n</i> (%)	Persistent long-COVID <i>n</i> = 139 <i>n</i> (%)	Ever vs never RR ^a (95% CI)	Persistent vs never RR ^a (95% CI)
Age, years					
< 50	309 (14.6)	110 (17.0)	24 (17.3)	ref	ref
50–64	1411 (66.7)	438 (67.7)	97 (69.8)	0.90 (0.75,1.07)	0.81 (0.53,1.25)
65 +	397 (18.8)	99 (15.3)	18 (12.9)	0.77 (0.61,0.98)	0.55 (0.30,1.01)
Sex					
Male	841 (39.7)	184 (28.4)	36 (25.9)	ref	ref
Female	1276 (60.3)	463 (71.6)	103 (74.1)	1.47 (1.26,1.71)	1.80 (1.25,2.60)
Educational level					
University	1183 (55.9)	322 (49.8)	53 (38.1)	ref	ref
Secondary	776 (36.7)	267 (41.3)	68 (48.9)	1.21 (1.05,1.40)	1.92 (1.35,2.72)
Primary or less	158 (7.5)	58 (9.0)	18 (12.9)	1.31 (1.03,1.67)	2.63 (1.57,4.41)
Deprivation index, mean (sd)	− 0.9 (0.7)	− 0.8 (0.7)	− 0.8 (0.7)	1.08 (0.98,1.19)	1.08 (0.86,1.36)
Smoking status					
Never smoker	916 (43.3)	256 (39.6)	65 (46.8)	ref	ref
Ex-smoker	916 (43.3)	304 (47.0)	61 (43.9)	1.13 (0.97,1.31)	0.87 (0.62,1.23)
Smoker	284 (13.4)	87 (13.4)	13 (9.4)	1.02 (0.82,1.26)	0.57 (0.32,1.03)
Alcohol intake					
No	454 (21.5)	166 (25.8)	47 (34.1)	ref	ref
Yes, occasionally	1360 (64.3)	391 (60.7)	76 (55.1)	0.90 (0.77,1.05)	0.66 (0.46,0.94)
Yes, every day	302 (14.3)	87 (13.5)	15 (10.9)	1.00 (0.79,1.27)	0.72 (0.40,1.30)
Physical activity					
Low	374 (17.9)	148 (23.1)	42 (31.3)	ref	ref
Moderate	993 (47.4)	289 (45.2)	53 (39.6)	0.79 (0.67,0.94)	0.49 (0.33,0.72)
High	728 (34.7)	203 (31.7)	39 (29.1)	0.80 (0.67,0.96)	0.53 (0.35,0.81)
Sleep hours					
≤ 5	170 (8.0)	77 (11.9)	22 (15.8)	1.31 (1.06,1.62)	1.72 (1.08,2.73)
6	600 (28.4)	178 (27.5)	34 (24.5)	1.00 (0.85,1.18)	0.91 (0.60,1.37)
7	900 (42.5)	264 (40.8)	56 (40.3)	ref	ref
8	401 (19.0)	97 (15.0)	18 (12.9)	0.84 (0.69,1.04)	0.70 (0.42,1.18)
≥ 9	45 (2.1)	31 (4.8)	9 (6.5)	1.74 (1.30,2.34)	2.62 (1.38,4.98)
BMI, mean (sd)	26.2 (4.3)	27.1 (4.9)	27.2 (4.8)	1.03 (1.02,1.04)	1.04 (1.01,1.07)
BMI ≥ 30 kg/m ²	355 (16.8)	158 (24.4)	38 (27.3)	1.39 (1.20,1.62)	1.66 (1.17,2.37)

There are small differences in the total number of participants for some variables due to missing values

RR relative risk, CI confidence intervals

^a Adjusted for age, sex, and educational level

Antibody responses

Among SARS-CoV-2 infected participants who underwent serological testing early in the pandemic (at baseline around June 2020) (*n* = 442), antibody responses were elevated in individuals who later developed long-COVID, in contrast to those who did not (Table 5 and Additional File 1: Fig. S2). This difference was evident across all five antigens examined, with a tendency for wider differences in all nucleocapsid antigens. Adjusting for COVID-19

severity reduced slightly these associations but all associations remained statistically significant.

Risk factors for persistent long-COVID

Risk factors for persistent long-COVID compared to no long-COVID are shown in Tables 2, 3, and 4 and are in the same direction overall but more pronounced, both for risk and protective factors, than those observed for all long-COVID cases. For example, those with severe/

Table 3 Self-rated health status and prior chronic diseases and their association (RR, (95% CI)) with long-COVID development and persistence in 2764 SARS-CoV-2-infected participants, COVICAT cohort 2020–2023

	Never long-COVID	Ever long-COVID	Persistent long-COVID	Ever vs. never	Persistent vs. never
	<i>n</i> = 2117 <i>n</i> (%)	<i>n</i> = 647 <i>n</i> (%)	<i>n</i> = 139 <i>n</i> (%)	RR ^a (95% CI)	RR ^a (95% CI)
Self-appraisal of health status					
Excellent/great	854 (40.3)	177 (27.4)	30 (21.6)	ref	ref
Fine	1040 (49.1)	313 (48.4)	64 (46.0)	1.33 (1.13,1.57)	1.63 (1.06,2.51)
Regular/bad	223 (10.5)	157 (24.3)	45 (32.4)	2.26 (1.89,2.71)	4.18 (2.68,6.50)
Chronic disease	671 (31.7)	299 (46.2)	78 (56.1)	1.58 (1.38,1.80)	2.50 (1.81,3.45)
Asthma or COPD	190 (9.0)	94 (14.5)	24 (17.3)	1.45 (1.21,1.73)	1.98 (1.32,2.98)
Hypertension	234 (11.1)	99 (15.3)	30 (21.6)	1.42 (1.18,1.70)	2.27 (1.56,3.32)
Heart disease	32 (1.5)	15 (2.3)	7 (5.0)	1.46 (0.96,2.24)	3.06 (1.52,6.17)
Digestive diseases	62 (2.9)	44 (6.8)	17 (12.2)	1.79 (1.41,2.25)	3.52 (2.27,5.45)
Joint problems/rheumatism	115 (5.4)	71 (11.0)	21 (15.1)	1.69 (1.39,2.06)	2.69 (1.75,4.11)
Autoimmune disease	52 (2.5)	27 (4.2)	7 (5.0)	1.40 (1.02,1.93)	1.85 (0.90,3.78)
Anxiety/depression	60 (2.8)	51 (7.9)	13 (9.4)	1.95 (1.58,2.42)	2.88 (1.71,4.86)

RR relative risk, CI confidence intervals

^a Adjusted for age, sex, and educational level

critical COVID-19 had a RR of 109.67 (95% CI 26.82–448.45) of reporting persistent long-COVID symptoms compared to those infected without long-COVID symptoms. A similar pattern was shown for the association of IgG levels in 2020 with persistent long-COVID (Table 5 and Additional File 1: Fig. S2).

Overall long-COVID predictors, multivariable model

A multivariable analysis including all factors that demonstrated statistically significant associations with long-COVID risks, while keeping age, sex, and educational level as fixed factors in the model (Table 6), identified BMI, asthma/COPD, anxiety/depression, digestive diseases, joint problems, COVID-19 severity, and pre-infection vaccination as the most robust predictors of long-COVID. Antibody responses were available only for a sub-sample and were not included in the multivariate model.

In an alternative analysis, we applied the false discovery rate correction (Additional File 1: Table S5). Results were very similar to those observed without multiple comparison corrections.

Discussion

In this population-based adult cohort, we observed (i) that 23% of infected individuals met our definition of long-COVID; (ii) three long-COVID subtypes; (iii) persistent long-COVID symptoms after two years in about half of the cases; (iv) female sex, COVID-19 severity, increased antibody response following primary infection, and prior chronic conditions including obesity as

risk factors; and (v) omicron infection, vaccination and to a lesser extent lifestyle, specifically physical activity and 6–8 h sleep, as preventive factors.

A challenge in understanding long-COVID remains the lack of identification of the source population of individuals diagnosed, and a standardized definition, despite the core symptoms showing consistency across studies [4, 24–27]. Additionally, reported prevalence rates are contingent on the duration of follow-up in each study. The cases examined in our study were drawn from a well-identified population-based cohort followed up at critical times during the pandemic (early 2020, 2021, and 2023). The long-term follow-up of our population allows robust estimates of the risk of long-COVID independent of the phase of the pandemic and description of the long-term course of the condition. While the participation rate declined over time, characteristics of individuals lost to follow-up were not markedly dissimilar to those remaining in the study regarding COVID-19 severity, minimizing potential biases.

We had the timing of occurrence of SARS-CoV-2 infection, exact dates for clinical diagnosis, and for vaccination, but we did not have exact dates for the occurrence of long-COVID symptoms and could therefore, only evaluate cumulative incidence at the times of contact.

We tracked the population through late spring 2023, capturing long-COVID cases with approximately two years of follow-up. Half of the cases indicated that their symptoms endured, underscoring the critical importance of sustained clinical surveillance and treatment for long-COVID patients well beyond the initial infection.

Table 4 COVID-19 severity, variants, and vaccination and their association with long-COVID development and persistence in 2764 SARS-CoV-2-infected participants of the COVICAT cohort, 2020–2023

	Never long-COVID <i>n</i> = 2117 <i>n</i> (%)	Ever long-COVID <i>n</i> = 647 <i>n</i> (%)	Persistent long-COVID <i>n</i> = 139 <i>n</i> (%)	Ever vs. never RR* (95% CI)	Persistent vs. never RR ^{*,†} (95% CI)
COVID-19 severity					
Asymptomatic	384 (18.1)	32 (4.9)	2 (1.4)	ref	ref
Mild/moderate	1712 (80.9)	564 (87.2)	115 (82.7)	3.10 (2.20,4.36)	11.67 (2.89,47.21)
Severe/critical	21 (1.0)	51 (7.9)	22 (15.8)	9.88 (6.88,14.18)	109.67 (26.82,448.45)
First infection					
Pre-omicron period	557 (26.8)	379 (58.6)	139 (100.0)	ref	
Omicron period	1525 (73.2)	268 (41.4)	0 (0.0)	0.36 (0.32,0.42)	
Vaccination pre-infection					
No	429 (21.0)	356 (55.0)	137 (98.6)	ref	ref
Yes	1617 (79.0)	291 (45.0)	2 (1.4)	0.33 (0.29,0.38)	0.01 (0.00,0.02)
Vaccination pre-infection, by amount					
No vaccine	429 (21.2)	356 (55.0)	137 (98.6)	ref	
1 dose	27 (1.3)	14 (2.2)	2 (1.4)	0.78 (0.52,1.19)	
2 doses	351 (17.4)	82 (12.7)	0 (0.0)	0.40 (0.33,0.50)	
3 doses	1170 (57.9)	193 (29.8)	0 (0.0)	0.31 (0.27,0.36)	
4 doses	42 (2.1)	2 (0.3)	0 (0.0)	0.10 (0.03,0.40)	
Vaccination within 3 months post-infection					
No	373 (88.6)	338 (94.9)	132 (96.4)	ref	ref
Yes	48 (11.4)	18 (5.1)	5 (3.6)	0.58 (0.39,0.86)	0.32 (0.09,1.19)
Type of first vaccine dose					
BioNTech/Pfizer	849 (50.3)	157 (54.0)	0 (0.0)	ref	
Moderna/Lonza	197 (11.7)	35 (12.0)	0 (0.0)	1.01 (0.72,1.42)	
Oxford/AstraZeneca	611 (36.2)	93 (32.0)	2 (100.0)	0.96 (0.74,1.26)	
J&J/Janssen	31 (1.8)	6 (2.1)	0 (0.0)	1.11 (0.53,2.30)	

RR relative risk, CI confidence intervals

There are small differences in the total number of participants for some variables due to missing values

Table 5 SARS-CoV-2 antibody levels in 2020 (pre-vaccination) and their association with long-COVID development and persistence in 442 SARS-CoV-2-infected participants and seropositive in 2020

SARS-CoV-2 antibody levels ^a	Never long-COVID <i>n</i> = 304 Mean (SD)	Ever long-COVID <i>n</i> = 138 Mean (SD)	Persistent long-COVID <i>n</i> = 52 Mean (SD)	Ever vs never		Persistent vs never	
				RR ^b (95% CI)	RR ^c (95% CI)	RR ^b (95% CI)	RR ^c (95% CI)
IgG NFL	4.2 (0.5)	4.5 (0.5)	4.5 (0.5)	2.14 (1.60,2.87)	1.77 (1.36,2.30)	3.15 (1.81,5.48)	2.52 (1.53,4.14)
IgG Nct	3.4 (0.5)	3.7 (0.6)	3.8 (0.6)	2.00 (1.58,2.54)	1.66 (1.31,2.09)	2.70 (1.78,4.08)	2.10 (1.39,3.18)
IgG RBD	3.1 (0.7)	3.6 (0.8)	3.7 (0.8)	1.67 (1.41,1.98)	1.45 (1.24,1.71)	2.13 (1.56,2.93)	1.83 (1.35,2.48)
IgG S	3.3 (0.8)	3.8 (0.8)	3.9 (0.8)	1.60 (1.34,1.91)	1.41 (1.20,1.66)	2.08 (1.47,2.93)	1.80 (1.31,2.48)
IgG S2	4.1 (0.6)	4.5 (0.6)	4.5 (0.6)	1.82 (1.44,2.31)	1.57 (1.27,1.96)	2.42 (1.55,3.76)	2.04 (1.36,3.07)

SD standard deviation; RR relative risk, CI confidence intervals

^a log₁₀ transformed^b Adjusted for age, sex, and educational level^c Adjusted for age, sex, educational level, and COVID-19 severity

Table 6 Multivariable model on the association between sociodemographic, lifestyle, clinical, and COVID-19-related factors and long-COVID development among 2665 SARS-CoV-2-infected participants of the COVICAT cohort, 2020–2023

	RR (95% CI)
Age, years	
< 50	ref
50–64	0.89 (0.75, 1.06)
65 +	0.82 (0.66, 1.03)
Sex	
Male	ref
Female	1.47 (1.27, 1.69)
Educational level	
University	ref
Secondary	1.09 (0.96, 1.25)
Primary or less	1.25 (1.01, 1.56)
BMI ≥ 30 kg/m ²	1.19 (1.03, 1.37)
Asthma or COPD	1.35 (1.14, 1.59)
Anxiety/depression	1.36 (1.09, 1.71)
Digestive diseases	1.33 (1.04, 1.70)
Joint problems, rheumatism	1.40 (1.14, 1.71)
COVID-19 severity	
Asymptomatic	ref
Mild/moderate	2.42 (1.74, 3.38)
Severe/critical	4.31 (2.99, 6.20)
Vaccination pre-infection, by amount	
No vaccine	ref
1 dose	0.81 (0.55, 1.21)
2 doses	0.44 (0.36, 0.54)
3 doses	0.35 (0.30, 0.41)
4 doses	0.10 (0.03, 0.38)

RR relative risk, CI confidence intervals, BMI body mass index

SARS-CoV-2 infection, COVID-19, and long-COVID are reported to increase the risk of several medical conditions including respiratory conditions, cardiac arrest, or type 2 diabetes, highlighting the necessity to understand the course of long-COVID symptoms over time [10].

Using hypothesis-free methods, we identified three subtypes with differing clinical presentation and prognostic ability, as the subtype severe multi-organ had a higher risk for long-COVID persistence. Our findings emphasize the need for comprehensive, longitudinal, and deep phenotyping studies to validate the clinical applicability of long-COVID subtyping.

While the highest risks were observed among individuals with severe COVID-19, nearly ten times as many cases of long-COVID emerged among patients with mild manifestations of the disease since they constitute the vast majority of COVID-19 patients. Our findings on disease severity were supported by strong correlations between 2020 IgG pre-vaccination antibody levels following initial

SARS-CoV-2 infection, which exhibit a close correlation with infection severity [15], and subsequent development of long-COVID. Nevertheless, after adjusting for COVID-19 severity, higher IgG were still associated with long-COVID, suggesting that the immune system mediates (e.g., through immune activation) or reflects (e.g., virus persistence) the underlying mechanisms implicated in long-COVID development. Also in line with these results and previous studies [7], we found a lower risk for long-COVID symptoms in participants infected with omicron variants compared to pre-omicron. However, we did not investigate the factors underlying this association, which may be attributed to the characteristics of the omicron variant, known to cause milder infections, as well as other factors, such as widespread vaccination. The discovery that pre-infection vaccination confers a protective effect on long-COVID, even when accounting for COVID-19 severity, suggests improved infection control and immune response modulation by vaccines.

We and others have shown that women are at higher risk of developing long-COVID [30–32], although they are at lower risk of developing severe infection. The potential underlying mechanisms are complex to discuss, and probably heterogeneous. Moreover, this female predominance aligns with broader patterns seen in immune-mediated diseases particularly auto-immune diseases, where women are generally more vulnerable [33]. More data would be needed to deepen into the potential factors involved.

We observed that lifestyle factors, namely physical activity, sleep, and obesity, were associated with long-COVID. These factors are potentially modifiable and thus crucial in preventing a chronic condition. The well-established impact of physical activity on the onset and prognosis of respiratory conditions should be taken into account in patients affected by long-COVID [34, 35]. Sleep duration exhibits a U-shaped relationship with health outcomes, highlighting its importance in preventing long-COVID [36]. Finally, obesity, a lifestyle disease that increases susceptibility to infections [37], remained in the multivariable model underscoring its role as a predictor for the development of long-COVID. We demonstrated that smoking is a risk factor specifically for omicron-related long-COVID. Earlier variants like the original strain and Delta caused more severe (lung) infections, potentially masking the long-term effects observed in smokers.

The influence of pre-existing chronic conditions, notably respiratory, digestive, joints diseases, anxiety, and depression, exerts a substantial role in long-COVID development and persistence. It is recognized that long COVID can also exacerbate pre-existing health conditions [4]. Asthma/COPD are likely

contributing to the increased risk through pre-existing lung damage as well as immune and inflammatory-related pathways. The role of mental health conditions [38, 39] may have been previously underestimated due to challenges in establishing temporal sequences. Immune dysregulation, increasingly recognized in mental health diseases including depression, could explain the higher risk of long-COVID [40].

Our study has several strengths and some limitations. Among the strengths are (i) the population-based cohort design, which captures a wide spectrum of long-COVID symptoms and incorporates novel exposures, offering advantages over studies based solely on clinical records; (ii) the comprehensive assessment of long-COVID symptoms, particularly valuable during the 2021 follow-up when the condition was still poorly characterized, together with the integration of electronic health record data and serological testing conducted in a substantial portion of the cohort; and (iii) the long follow-up period that included pre-pandemic health data, enabling the evaluation of selection processes within the population. However, the study also has limitations: (i) losses to follow-up, which may have introduced selection bias. Nevertheless, we demonstrated that prior health status, including the severity of COVID-19, was not significantly related to the probability of follow-up, reducing the likelihood of bias toward more severe cases; (ii) the limited frequency of follow-up contacts during the pandemic. While more frequent contacts would have been advantageous, they were not feasible due to the population-based, non-hospitalized nature of the study. Additionally, EHRs often inadequately capture the diverse spectrum of long-COVID symptoms, limiting their utility in providing timely additional information for most cases including treatment; (iii) the timing of the first infection was not possible to document for asymptomatic individuals with serological evidence of infection. Nonetheless, this limitation did not affect the validity of our analyses, as we evaluated broad pandemic periods (e.g., pre-Omicron vs. Omicron) that were robustly defined; (iv) the lack of appropriate data of re-infections, as SARS-CoV-2 testing was not widespread at the population level at the end of the study period and serology is not effective for detecting re-infections due to prior immunity; and (v) similar to other studies on long COVID [41], information on long COVID symptoms was requested only among those infected. This does not affect inferences on the risk of long COVID among the infected, although may limit the conduct of other analyses such as the specificity of long-term symptoms that are attributable to SARS-CoV-2 infection.

Conclusions

In conclusion, our study reveals a significant risk of long-COVID after SARS-CoV-2 infection and found that symptoms persist for over two years in about half of cases. The risk of long-COVID was influenced by various factors, with the severity of the initial SARS-CoV-2 infection, COVID-19 vaccination, and prior chronic conditions playing predominant roles. International collaboration in population-based studies will be essential to assess the generalizability of these findings across diverse populations.

Abbreviations

SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
COVID-19	Coronavirus disease 2019
WHO	World Health Organization
IgM	Immunoglobulin M
IgG	Immunoglobulin G
IgA	Immunoglobulin A
S	Spike full protein
S2	S2 fragment
RBD	Receptor-binding domain
NFL	Nucleocapsid full protein
NCt	Nucleocapsid C-terminal region
COPD	Chronic obstructive pulmonary disease
BMI	Body mass index
MFI	Median fluorescence intensity

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-025-03974-7>.

Additional file 1: Table S1. Characteristics of responders versus non-responders; Table S2-Questions in 2021 and 2023 surveys to evaluate long COVID symptoms; Table S3- Frequency of reported long term symptoms total and by sex; Table S4- Association [RR, (95% CI)] of risk factors with long-COVID development stratified by variant of infection ; Table S5- P-values for risk factors evaluated without corrections for multiple comparisons, and p-values with Benjamini-Hochberg correction. Figure S1-S2. Fig S1- Long term symptoms by variant of first infection. Fig S2- SARS-CoV-2 IgG levels in 2020 against five SARS-CoV-2 antigens by ever long-COVID status.

Acknowledgements

We thank the volunteers who participate in the GCAT cohort study. We thank all the workers in different facilities of the Blood and Tissue Bank from Catalonia (BST) and especially Dr. Joan Grifols. This study was carried out using anonymized data provided by the Catalan Agency for Quality and Health Assessment, within the framework of the PADRIIS Program. The authors of this study would like to acknowledge all GCAT project investigators who contributed to the generation of the GCAT data, especially to Alba Blasco, Beatriz Cortés and Anna Carreras and to ISGlobal investigators who generated the SARS-CoV-2 antibody data, particularly Alfons Jiménez, Ruth Aguilar, Marta Vidal, Luis Izquierdo; and to Pere Santamaria for antigens. A full list of the investigators is available from <http://www.genomesforlife.com/>.

We gratefully acknowledge the authors, originating and submitting laboratories of the sequences from GISAID's EpiCov™ Database. All submitters of data may be contacted directly via <http://www.gisaid.org/>.

Authors' contributions

All authors satisfy the criteria for authorship. MK1, MK2, GM, CD, RdC, NP and JGA designed the study; AE, SI, MBdB and LDO did the data editing and statistical analysis; SIG, GCV, NB, EAN, GM and CD contributed in data acquisition, MK1, MK2, AE, GM, CD, XF, NP, RdC and JGA drafted the manuscript and all authors contributed to interpretation of the work. All authors reviewed the

manuscript critically for important intellectual content, gave final approval of the version to be published, and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Authors' Twitter handles

Twitter handles: @MalarialImmuno; @gmoncu (Gemma Moncunill); @judith-garciaaym (Judith Garcia-Aymerich).

Funding

We acknowledge support from the Horizon Europe END-VOC (grant agreement no. 101046314), Spanish Ministry of Science & Innovation (PID2019-110810RB-I00 grant); the Spanish State Research Agency and Ministry of Science and Innovation through the "Centro de Excelencia Severo Ochoa 2019–2023" Program (CEX2018-000806-S), the Instituto de Salud Carlos III (PI17/01555, PI18/01512), the La MaratóTV3 Foundation (167/C/2021), the Generalitat de Catalunya through the CERCA Program and the Fundació Privada Daniel Bravo Andreu; Xavier Farré is supported by VEIS project (001-P-001647, co-funded by European Regional Development Fund (ERDF). G.M. is supported by RYC 2020–029886-I/AEI/<https://doi.org/10.13039/501100011033>, co-funded by the European Social Fund (ESF). We acknowledge support from the grant CEX2023-0001290-S funded by MCIN/AEI/ <https://doi.org/10.13039/501100011033>, and support from the Generalitat de Catalunya through the CERCA Program.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. Protocol information will be available on reasonable request.

Declarations

Ethics approval and consent to participate

All participants contacted had consented in the past to be re-contacted and had provided informed consent. Ethical approval 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).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹ISGlobal, Barcelona, Spain. ²CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain. ³Universitat Pompeu Fabra (UPF), Barcelona, Spain. ⁴IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain. ⁵Genomes for Life–GCAT Lab, CORE Program. Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain. ⁶Grup de Recerca en Impacte de Les Malalties Cròniques i Les Seves Trajectòries (GRIMTra) (IGTP), Badalona, Spain. ⁷London School of Hygiene and Tropical Medicine, London, UK. ⁸Banc de Sang i Teixits (BST), Barcelona, Spain. ⁹CIBER Enfermedades Infecciosas (CIBERINFEC), Barcelona, Spain.

Received: 18 September 2024 Accepted: 27 February 2025

Published: 14 March 2025

References

- Høeg TB, Ladhani S, Prasad V. How methodological pitfalls have created widespread misunderstanding about long COVID. *BMJ EBM*. 2023;29(3):142–6.
- Global Burden of Disease Long COVID Collaborators, Wulf Hanson S, Abbafati C, Aerts JG, Al-Aly Z, Ashbaugh C, et al. Estimated global proportions of individuals with persistent fatigue, cognitive, and respiratory symptom clusters following symptomatic COVID-19 in 2020 and 2021. *JAMA*. 2022;328(16):1604.
- Thaweethai T, Jolley SE, Karlson EW, Levitan EB, Levy B, McComsey GA, et al. Development of a Definition of Postacute Sequelae of SARS-CoV-2 Infection. *JAMA*. 2023;329(22):1934.
- Committee on Examining the Working Definition for Long COVID, Board on Health Sciences Policy, Board on Global Health, Health and Medicine Division, National Academies of Sciences, Engineering, and Medicine. A Long COVID Definition: A Chronic, Systemic Disease State with Profound Consequences. Fineberg HV, Brown L, Worku T, Goldowitz I, editors. Washington, D.C.: National Academies Press; 2024. Available from: <https://www.nap.edu/catalog/27768>. Cited 2024 Dec 26.
- Ballering AV, Van Zon SKR, Olde Hartman TC, Rosmalen JGM. Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. *The Lancet*. 2022;400(10350):452–61.
- Thompson EJ, Williams DM, Walker AJ, Mitchell RE, Niedzwiedz CL, Yang TC, et al. Long COVID burden and risk factors in 10 UK longitudinal studies and electronic health records. *Nat Commun*. 2022;13(1):3528.
- Xie Y, Choi T, Al-Aly Z. Postacute sequelae of SARS-CoV-2 Infection in the Pre-Delta, Delta, and Omicron Eras. *N Engl J Med*. 2024;391(6):515–25.
- Hadley E, Yoo YJ, Patel S, Zhou A, Laraway B, Wong R, et al. Insights from an N3C RECOVER EHR-based cohort study characterizing SARS-CoV-2 reinfections and Long COVID. *Commun Med*. 2024;4(1):129.
- Bowe B, Xie Y, Al-Aly Z. Acute and postacute sequelae associated with SARS-CoV-2 reinfection. *Nat Med*. 2022Nov;28(11):2398–405.
- Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol*. 2023;21(3):133–46.
- Cervia-Hasler C, Brüningk SC, Hoch T, Fan B, Muzio G, Thompson RC, et al. Persistent complement dysregulation with signs of thromboinflammation in active Long Covid. *Science*. 2024;383(6680):eadg7942.
- Sherif ZA, Gomez CR, Connors TJ, Henrich TJ, Reeves WB. RECOVER Mechanistic Pathway Task Force. Pathogenic mechanisms of post-acute sequelae of SARS-CoV-2 infection (PASC). *eLife*. 2023;12:e86002.
- Reme BA, Gjesvik J, Magnusson K. Predictors of the post-COVID condition following mild SARS-CoV-2 infection. *Nat Commun*. 2023;14(1):5839.
- Subramanian A, Nirantharakumar K, Hughes S, Myles P, Williams T, Gokhale KM, et al. Symptoms and risk factors for long COVID in non-hospitalized adults. *Nat Med*. 2022;28(8):1706–14.
- Karachaliou M, Moncunill G, Espinosa A, Castaño-Vinyals G, Jiménez A, Vidal M, et al. Infection induced SARS-CoV-2 seroprevalence and heterogeneity of antibody responses in a general population cohort study in Catalonia Spain. *Sci Rep*. 2021;11(1):21571.
- Karachaliou M, Moncunill G, Espinosa A, Castaño-Vinyals G, Rubio R, Vidal M, et al. SARS-CoV-2 infection, vaccination, and antibody response trajectories in adults: a cohort study in Catalonia. *BMC Med*. 2022;20(1):347.
- Kogevinas M, Castaño-Vinyals G, Karachaliou M, Espinosa A, de Cid R, Garcia-Aymerich J, et al. 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*. 2021;129(11):117003.
- Delgado-Ortiz L, Carsin AE, Merino J, Cobo I, Koch S, Goldberg X, et al. Changes in Population Health-Related Behaviors During a COVID-19 Surge: A Natural Experiment. *Ann Behav Med*. 2023;57(3):216–26.
- Obón-Santacana M, Vilardell M, Carreras A, Duran X, Velasco J, Galván-Femenía I, et al. GCAT[Genomes for life: a prospective cohort study of the genomes of Catalonia. *BMJ Open*. 2018;8(3):e018324.
- Rinaldi E, Stellmach C, Rajkumar NMR, Carocchia N, Dellacasa C, Giannella M, et al. Harmonization and standardization of data for a pan-European cohort on SARS-CoV-2 pandemic. *NPJ Digit Med*. 2022;5(1):75.
- Dobaño C, Santano R, Jiménez A, Vidal M, Chi J, Rodrigo Melero N, et al. Immunogenicity and crossreactivity of antibodies to the nucleocapsid protein of SARS-CoV-2: utility and limitations in seroprevalence and immunity studies. *Transl Res*. 2021;232:60–74.
- Dobaño C, Vidal M, Santano R, Jiménez A, Chi J, Barrios D, et al. Highly sensitive and specific multiplex antibody assays to quantify immunoglobulins M, A, and G against SARS-CoV-2 antigens. *J Clin Microbiol*. 2021;59(2):e01731–20.
- WHO. WHO COVID-19 Clinical management: living guidance. Geneva: WHO; 2020 May. Report No.: WHO/2019-nCoV/clinical/2021.2.
- Soriano JB, Murthy J, Marshall JC, Relan P, Diaz JV. WHO clinical case definition working group on post-COVID-19 condition. A clinical case

- definition of post-COVID-19 condition by a Delphi consensus. *Lancet Infect Dis.* 2022;22(4):e102–7.
25. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, Sepulveda R, Rebolledo PA, Cuapio A, et al. More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. *Sci Rep.* 2021;11(1):16144.
 26. Havervall S, Rosell A, Phillipson M, Mangsbo SM, Nilsson P, Hober S, et al. Symptoms and Functional Impairment Assessed 8 Months After Mild COVID-19 Among Health Care Workers. *JAMA.* 2021;325(19):2015–6.
 27. Carfi A, Bernabei R, Landi F. for the Gemelli Against COVID-19 post-acute care study group. Persistent symptoms in patients after acute COVID-19. *JAMA.* 2020;324(6):603.
 28. Duque I, Domínguez-Berjón MF, Cebrecos A, Prieto-Salceda MD, Esnaola S, Calvo Sánchez M, et al. Deprivation index by enumeration district in Spain, 2011. *Gac Sanit.* 2021;35(2):113–22.
 29. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Stat Methodol.* 1995;57(1):289–300.
 30. Bai F, Tomasoni D, Falcinella C, Barbanotti D, Castoldi R, Mulè G, et al. Female gender is associated with long COVID syndrome: a prospective cohort study. *Clin Microbiol Infect.* 2022;28(4):611.e9–611.e16.
 31. Cohen J, van der Meulen RY. An intersectional analysis of long COVID prevalence. *Int J Equity Health.* 2023;22(1):261.
 32. Pérez Catalán I, Roig Martí C, Fabra Juana S, Domínguez Bajo E, Herero Rodríguez G, Segura Fábrega A, et al. One-year quality of life among post-hospitalization COVID-19 patients. *Front Public Health.* 2023;6(11):1236527.
 33. Dou DR, Zhao Y, Belk JA, Zhao Y, Casey KM, Chen DC, et al. Xist ribonucleoproteins promote female sex-biased autoimmunity. *Cell.* 2024;187(3):733–749.e16.
 34. Carsin AE, Keidel D, Fuertes E, Imboden M, Weyler J, Nowak D, et al. Regular Physical Activity Levels and Incidence of Restrictive Spirometry Pattern: A Longitudinal Analysis of 2 Population-Based Cohorts. *Am J Epidemiol.* 2020;189(12):1521–8.
 35. Vaes AW, Garcia-Aymerich J, Marott JL, Benet M, Groenen MTJ, Schnohr P, et al. Changes in physical activity and all-cause mortality in COPD. *Eur Respir J.* 2014;44(5):1199–209.
 36. Chaput JP, Dutil C, Featherstone R, Ross R, Giangregorio L, Saunders TJ, et al. Sleep duration and health in adults: an overview of systematic reviews. *Appl Physiol Nutr Metab.* 2020;45(10 (Suppl. 2)):S218–31.
 37. Pugliese G, Liccardi A, Graziadio C, Barrea L, Muscogiuri G, Colao A. Obesity and infectious diseases: pathophysiology and epidemiology of a double pandemic condition. *Int J Obes (Lond).* 2022;46(3):449–65.
 38. Wang S, Quan L, Chavarro JE, Slopen N, Kubzansky LD, Koenen KC, et al. Associations of Depression, Anxiety, Worry, Perceived Stress, and Loneliness Prior to Infection With Risk of Post-COVID-19 Conditions. *JAMA Psychiat.* 2022;79(11):1081–91.
 39. Pérez Catalán I, Roig Martí C, Folgado Escudero S, Segura Fábrega A, Varea Villanueva M, Fabra Juana S, et al. Presence of COVID-19 self-reported symptoms at 12 months in patients discharged from hospital in 2020–2021: a Spanish cross-sectional study. *Sci Rep.* 2024;14(1):26575.
 40. Dantzer R. Neuroimmune Interactions: From the Brain to the Immune System and Vice Versa. *Physiol Rev.* 2018;98(1):477–504.
 41. Hou Y, Gu T, Ni Z, Shi X, Ranney ML, Mukherjee B. Global Prevalence of Long COVID, its Subtypes and Risk factors: An Updated Systematic Review and Meta-Analysis. 2025. Available from: <https://medrxiv.org/lookup/doi/10.1101/2025.01.01.24319384>. Cited 2025 Jan 31.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.