



OPEN Identifying risk factors and predicting long COVID in a Spanish cohort

Antonio Guillén-Teruel¹, Jose L. Mellina-Andreu¹, Gabriel Reina², Enrique González-Billalabeitia³, Ramón Rodríguez-Iborra⁴, José Palma¹, Juan A. Botía¹ & Alejandro Cisterna-García¹✉

Many studies have investigated symptoms, comorbidities, demographic factors, and vaccine effects in relation to long COVID (LC-19) across global populations. However, a number of these studies have shortcomings, such as inadequate LC-19 categorisation, lack of sex disaggregation, or a narrow focus on certain risk factors like symptoms or comorbidities alone. We address these gaps by investigating the demographic factors, comorbidities, and symptoms present during the acute phase of primary COVID-19 infection among patients with LC-19 and comparing them to typical non-Long COVID-19 patients. Additionally, we assess the impact of COVID-19 vaccination on these patients. Drawing on data from the Regional Health System of the Region of Murcia in southeastern Spain, our analysis includes comprehensive information from clinical and hospitalisation records, symptoms, and vaccination details of over 675126 patients across 10 hospitals. We calculated age and sex-adjusted odds ratios (AOR) to identify protective and risk factors for LC-19. Our findings reveal distinct symptomatology, comorbidity patterns, and demographic characteristics among patients with LC-19 versus those with typical non-Long COVID-19. Factors such as age, female sex (AOR = 1.39, adjusted $p < 0.001$), and symptoms like chest pain (AOR > 1.55, adjusted $p < 0.001$) or hyposmia (AOR > 1.5, adjusted $p < 0.001$) significantly increase the risk of developing LC-19. However, vaccination demonstrates a strong protective effect, with vaccinated individuals having a markedly lower risk (AOR = 0.10, adjusted $p < 0.001$), highlighting the importance of vaccination in reducing LC-19 susceptibility. Interestingly, symptoms and comorbidities show no significant differences when disaggregated by type of LC-19 patient. Vaccination before infection is the most important factor and notably decreases the likelihood of long COVID. Particularly, mRNA vaccines offer more protection against developing LC-19 than viral vector-based vaccines (AOR = 0.48). Additionally, we have developed a model to predict LC-19 that incorporates all studied risk factors, achieving a balanced accuracy of 73% and ROC-AUC of 0.80. This model is available as a free online LC-19 calculator, accessible at ([LC-19 Calculator](#)).

Keywords COVID-19, SARS-CoV-2, Vaccines, Long COVID, LC-19

The virus accountable for Coronavirus Disease 2019 (COVID-19), namely the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a highly transmissible and pathogenic betacoronavirus that surfaced in Wuhan, China, in late 2019¹. According to the World Health Organization, there have been more than 776 million reported cases of COVID-19 and over 7 million deaths caused by SARS-CoV-2 as of October 2024². The severity, symptoms, and clinical progression of COVID-19 vary widely among individuals, spanning from asymptomatic cases to fatal outcomes^{3,4}. Moreover, the duration and severity of symptoms can vary significantly, with some individuals experiencing only brief symptoms lasting days or weeks, while others develop persistent symptoms^{5,6}. However, regardless of the duration or severity of the acute phase of COVID-19, all patients remain at risk of developing Long COVID.

Long COVID-19 (LC-19) represents a highly diverse condition encompassing symptoms that arise following a SARS-CoV-2 infection, with manifestations not attributable to any other known condition^{7,8}. In other words, LC-19 is usually diagnosed by a process of elimination of other potential diseases. Given the variability of LC-19,

¹Department of Information and Communication Engineering, University of Murcia, Murcia 30100, Spain. ²Servicio de Microbiología, Clínica, Universidad de Navarra, Instituto de Investigación Sanitaria de Navarra (IdiSNA), Pamplona, Navarra, Spain. ³Department of Medical Oncology, Hospital Universitario, 12 de Octubre, Madrid, Spain. ⁴Subdirección General de Tecnologías de la Información, Servicio Murciano de Salud, Murcia, Spain. ✉email: alejandro.cisterna@um.es

various definitions exist regarding the onset or duration of symptoms. However, the three most widely adopted subtypes of LC-19 are as follows: 1) Acute post-COVID symptoms emerge between week 5 and week 12 after the initial onset of symptoms; 2) Long post-COVID symptoms manifest from week 12 to week 24 following onset; and 3) Persistent post-COVID symptoms are characterised by lasting more than 24 weeks after the initial onset⁹.

The risk of developing LC-19 is influenced by various factors, including age, sex, comorbidities and vaccination status^{4,10–13}. For instance, Bai et al.¹⁴ determined that factors associated with an elevated risk of developing 'long COVID' include being female, advanced age, and active smoking. Interestingly, they found that the severity of the initial illness did not significantly contribute to the risk. Furthermore, Sylvester et al.¹⁵ conducted literature reviews and observed sex-disaggregated differences in the sequelae of COVID-19 and long COVID syndrome. They also pointed out the scarcity of studies reporting sex-disaggregated data on COVID-19, emphasising the urgent need for more research and reporting that takes sex-based differences into account.

Another important factor that should be taken into account is vaccination status. Several studies have been conducted to assess the impact of vaccines on the risk of LC-19^{16–21}. Logically, these studies have varied objectives and different study designs. For instance, Mohr et al.¹⁶ concluded that receipt of two doses of a COVID-19 mRNA was associated with a decreased prevalence of COVID-like symptoms at 6 weeks. Additionally, their research identified which COVID-19-like symptoms were more prevalent among unvaccinated versus vaccinated individuals. Notarte et al.¹⁷ conducted a systematic review finding evidence that vaccination before SARS-CoV-2 infection could reduce the risk of subsequent LC-19. In addition, in the context of LC-19 and vaccination status, Ayoubkhani et al.¹⁸ suggest that vaccination against COVID-19 reduces the likelihood of experiencing long COVID symptoms, however, they also emphasise the need for longer follow-up studies. In line with Català et al.¹⁹ conclusions that COVID-19 vaccination has been shown to significantly lower the incidence of long COVID symptoms, underscoring the critical role of vaccination in preventing the continuation of COVID-19 symptoms, especially among adults.

In this study, we aim to identify the factors that differentiate patients who develop LC-19 from those who do not. We examine demographic factors, comorbidities, and symptoms present during the acute phase to understand what sets individuals with LC-19 apart from other non-Long COVID-19 patients. Additionally, we examined the impact of COVID-19 vaccination (including the number of doses, administration schedules, and vaccine types) on the risk of developing LC-19. Finally, we develop a model to predict LC-19 by considering demographic data, pre-existing comorbidities, symptoms at initial diagnosis, and vaccination records before infection. Utilising data from the Regional Health System of the Region of Murcia in southeastern Spain, our study conducts a regional investigation by gathering information from 10 different hospitals. The dataset encompasses details from clinical records (including comorbidities), hospitalisation records, symptoms, and administered vaccines for over 675126 COVID-19 patients.

Methods

Study design, setting, and participants

In this study, we followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist for reporting²². The study protocol was approved by the Bioethics Committee of the Universidad de Murcia. Anonymized individual-level data from all patients diagnosed with COVID-19 within the Public Health Care System, collected in a prospective Regional Registry, were provided by the Murcia Health Care Information Department. A waiver for the need of informed consent was obtained in accordance with the General Data Protection Regulation (GDPR).

We conducted a retrospective cohort analysis using a population selected from a database provided by the Servicio Murciano de Salud (SMS). This database includes patients who tested positive for COVID-19 between January 4, 2020, and May 3, 2022, across 10 hospitals in the Region of Murcia, with both LC-19 and non-Long COVID-19 cases. Diagnosis of COVID-19 was confirmed by antigen testing or real-time reverse transcription-polymerase chain reaction (RT-PCR) assays using nasopharyngeal swab samples. RT-PCRs were performed in hospitals by clinicians and technicians, while antigen tests were conducted by nurses and clinicians in primary healthcare settings. The first inclusion criterion was being part of the COVID-19 database and having records related to symptoms, comorbidities, or vaccination status.

According to the Centers for Disease Control and Prevention (CDC)²³, a set of symptoms are more commonly associated with a case of LC-19, which we will refer to as LC-19 symptoms. The available LC-19 symptoms in our database include chest pain, muscle pain, cough, sleep problems, dyspnea, dizziness, hyposmia, hypogeusia, fever, headache, rash and tiredness. The symptoms included in our study were recorded, evaluated, and diagnosed by clinicians in primary healthcare settings or during hospitalisation. Using this information we distinguished LC-19 three subtypes of patients⁹ from non-Long COVID-19 patients following an operational definition. The time at which each patient symptom was recorded, evaluated, diagnosed or reported was considered the moment of appearance of the symptoms.

To identify non-Long COVID-19 patients, we selected individuals from the aforementioned database whose symptoms resolved within 28 days of testing positive for COVID-19. These patients' medical records were reviewed to confirm that their symptoms or reasons for follow-up appointments were not consistent with LC-19 symptoms, or that they had been diagnosed with other conditions. Additionally, patients who did not return for further medical attention at primary care or emergency services after their initial recovery were also classified as non-Long COVID-19.

For all patients, we assessed whether they exhibited LC-19 symptoms at least 28 days after their initial COVID-19 infection to identify those with LC-19. Acute post-COVID patients were defined as individuals whose LC-19 symptoms persisted for more than 28 days but less than 12 weeks after testing positive. Long post-COVID patients were those whose symptoms lasted between 12 and 24 weeks, while persistent post-COVID patients had symptoms that extended beyond 24 weeks. We followed the same process of reviewing medical

records and appointments for these patients as we did for non-Long COVID-19 patients. All subtypes of LC-19 patients were considered part of the broader LC-19 group.

The study also considered the vaccination status of the population from December 21, 2020, to May 3, 2022. Available data included specific vaccine brands, such as the mRNA vaccines BNT162b2 (Pfizer/BioNTech) and mRNA-1273 (Moderna), as well as the non-replicating viral vector vaccines ChAdOx1 (Oxford/AstraZeneca) and Ad26.COV2.S (Janssen). For a vaccination schedule to be validated, it had to be recorded at least 21 days prior to the patient's COVID-19 diagnosis.

After identifying each patient category, we assessed risk factors for developing LC-19, including symptoms at the time of the positive test, pre-existing comorbidities, sociodemographic factors, and vaccination status. These characteristics were compared between non-Long COVID-19 patients, LC-19 patients, and across the subtypes of LC-19 patients.

Ethics approval and consent to participate

All methods were performed following the relevant guidelines and regulations. The information of interest for the study has been obtained from the files of medical records of the Murcia Health Service in Spain, without the consent of the holders of the data. The need for informed consent was waived by the Bioethics Committee of Universidad de Murcia (See Ethics approval and consent to participate section in the [Supplementary Material](#)).

Variables and data source

The data used for this study is similar to that explained in the data description and preprocessing section of Cisterna et al. work²⁴. However, there is a new database called the 'vaccination database' that contains information on all patients in the Region of Murcia (Spain) during the period under review in this study previously specified. We examined demographic factors, comorbidities, symptoms, hospitalization status, and vaccine information (including the number of doses, administration schedules, and vaccine types) recorded in the COVID-19 database provided by the Servicio Murciano de Salud (SMS). Data regarding medication and length of stay in each hospital department were available, they were excluded from the analysis as they were deemed to be outside the scope of the study and would have significantly reduced the number of patients included due to low patient availability.

Regarding demographic factors, we have information about biological sex information (male or female) and the age of the patient (discrete quantitative variable). No other sociodemographic information was extracted or used in the analysis.

Regarding comorbidities, we have information from previous comorbidities associated with each patient. These comorbidities were diagnosed before COVID-19 by a clinician. We have information about diabetes mellitus, heart failure, chronic obstructive pulmonary disease, arterial hypertension, depression, ischemic cardiomyopathy, stroke, renal insufficiency, cirrhosis, osteoporosis, osteoarthritis, arthritis, obesity and asthma. All of these comorbidities are dichotomous qualitative variables. Additionally, we have information about the number of chronic diseases and the number of systems affected, which are discrete quantitative variables.

Regarding symptoms, we have information from symptoms at the initial diagnosis of the COVID-19 infection and up to 28 days after the positive test for COVID-19. The symptoms included in our study were recorded, evaluated, and diagnosed by clinicians in primary healthcare settings or during hospitalisation. We have information about chest pain, muscle pain, cough, lack of appetite, sleep problems, nasal congestion, dyspnea, rhinorrhoea, low grade fever, dizziness, hyposmia, hypogeusia, headache, tiredness, eye symptoms, expectorate, fever, sore throat, stomach pain, vomit, nasal discharge, chills, rash and malaise. All of these symptoms are dichotomous qualitative variables.

Furthermore, we have information regarding the severity of patients, as indicated by whether the patient has been hospitalised and, if so, the date of hospitalisation. This information is recorded by a clinician at the time the patient enters a hospital area.

Regarding vaccines status, we have information about the dosage, the brand of the vaccine (Astra Zeneca, Janssen, Moderna or Pfizer) and the schedule, which is recorded concurrently with the administration of the dose to each individual.

Bias

To ensure data quality and minimize bias, we designed a rigorous selection process to include individuals with comprehensive descriptions and complete data for all relevant features. The large sample size, derived from a retrospective cohort analysis using a health database, allowed for the exclusion of patients lacking essential information, such as comorbidities or demographic data. Cases with inconclusive or implausible records, such as two vaccine doses administered on the same day, negative age, or undetermined biological sex, were removed to maximize data accuracy and completeness. No NA values were recorded in the database. Therefore, any patient with at least one record in each of the symptom, comorbidity, and vaccine datasets, without any erroneous data, was considered part of the complete sample.

We applied methodologies to mitigate potential confounders and address timing-related variables. These adjustments were based on the timing of patients' positive COVID-19 test results and the onset of symptoms. We considered the time points when symptoms were first recorded, evaluated, diagnosed, or reported to clinicians to ensure accurate temporal alignment in the analysis.

A potential source of bias was sex, as prior studies suggest females may be at a higher risk of developing LC-19^{14,25–29}. To address this, we adjusted for sex in our analyses and performed sex-disaggregated assessments to reduce the influence of this bias.

Another potential bias is the lack of detail regarding the progression of symptoms over time. We only considered the presence or absence of symptoms within the first 28 days after a positive COVID-19 test, without

accounting for the specific days on which symptoms appeared. We acknowledge that this limitation may affect the accuracy of our model in capturing the temporal dynamics of symptom development.

One more possible source of bias is the time interval between vaccination and the development of immunity. To ensure valid immunization data, we only considered vaccination schedules where the final dose was administered at least 21 days prior to the patient's COVID-19 diagnosis. This criterion was set to account for the time required for the immune response to develop fully, thereby reducing the risk of misclassifying individuals who had not yet achieved full immunity.

Statistical analyses

Continuous data are presented as the median with interquartile range (IQR), while categorical data are expressed in percentages (%). Odds ratios (OR) along with 95% confidence intervals (CIs) were also employed. Following a pipeline to detect erroneous data and quantify them, we treated those values as NAs. Given their low proportion (2.54% of the total patients), we excluded these patients from the study, opting for a complete-case analysis. In various analyses, the ORs were adjusted for sex and age (as Adjusted Odds Ratios, AOR), and specifically for all comorbidities and symptoms within the respective analyses of comorbidities and symptomatology. Statistical analyses were conducted using R software (version 4.3.1), with a significance threshold for p-values set at 0.05. The Mann-Whitney U test is employed for the purpose of analysing differences in continuous data between two sets of patients. Differences between groups or conditions were evaluated by computing AOR, the associated 95% CI and p-value using the 'logistic.display' function of the R package *epiDisplay*³⁰ (version 3.5.0.2) on a generalised linear model (GLM), trained through the 'glm' function using R package *caret*³¹ (version 6.0-94), employing logistic regression. This approach employs logistic regression to predict a binary outcome from a set of predictors. The 'logistic.display' function is used to determine the weights associated with each variable, which in turn allows the AOR [95% CI] and p-values for all predictors to be calculated. (See Analysis performed to obtain the AORs section in the [Supplementary Material](#)).

AORs were also employed to examine differences among various vaccination schedules in the cohort, assessing whether specific schedules offer greater or lesser protection compared to others. We generated a heatmap that allows for pairwise comparisons of vaccination schedules, enabling the identification of statistically significant differences. We used p-values with Bonferroni correction to denote these differences and indicated which schedules are statistically significant (*) more protective or less protective. Additionally, an OR study was conducted to assess the comparative efficacy of mRNA-based vaccines and viral vector-based ones for the prevention of LC-19. This analysis was adjusted by age, sex, and the number of doses received. A further analysis was conducted to determine whether a combination of an mRNA-based vaccine and a COVID-19 viral vector-based vaccine is more protective than two separate vaccines for suffering LC-19. Two further analyses were conducted to ascertain whether a combination of one COVID-19 viral vector-based and one mRNA-based vaccine is more efficacious than two mRNA-based vaccines or two COVID-19 viral vector-based vaccines for suffering LC-19, with both analyses adjusted for sex and age.

LC-19 probability predictive model

A calculator capable of estimating the probability of developing LC-19 using machine learning (ML) techniques was developed. The predictive model included the following factors: age, sex, vaccination schedule, comorbidities and symptoms at the initial diagnosis of the COVID-19 infection and up to 28 days after the positive test for COVID-19. Vaccination schedules recorded prior to 21 days before the COVID-19 positive and previous comorbidities were considered for training the model. The data utilised for creating this tool include patients with all the factors previously mentioned available ($n = 289367$), and it was divided into two sets with the original LC-19 patients distribution: 70% for training and 30% for validation. As only a few patients suffer from LC-19, the dataset is highly imbalanced. Two techniques were employed to address this imbalance: IPID²⁴, a boosting technique that facilitates the training of ML models on balanced subdatasets, and SMOTE³², a synthetic oversampling technique that balances the original dataset. Both techniques were employed in conjunction with logistic regression models. The results were evaluated using the balanced accuracy and the area under the receiver operating characteristic curve (ROC-AUC).

Results

In the following subsection, we report the number and type of participants (see "[Participants](#)" section). Additionally, we present the characteristics studied and the preprocessing of the data (see "[Data description and preprocessing](#)" section). Subsequently, we explore demographic factors, comorbidities, and symptoms observed during the acute phase of primary infection in individuals with LC-19 and non-Long COVID-19 (see "[Analysis of sociodemographic factors, comorbidities and symptoms between LC-19 patients and non-Long COVID-19 patients](#)" section) and the effect of the severity of infection ("[Impact of COVID-19 hospitalisation on the development of LC-19](#)" section). Furthermore, we assess the influence of COVID-19 vaccination (see "[Effects of vaccination on LC-19](#)" section), including dosage, administration schedules, and types of vaccines, on LC-19 prevalence. Finally, we develop a model to predict the probabilities of developing LC-19 after a COVID-19 infection by considering demographic data, pre-existing comorbidities, symptoms at initial COVID-19 diagnosis, and vaccination schedules recorded prior to 21 days before the COVID-19 positive (see "[LC-19 calculator tool](#)" section).

Participants

A total of 675126 patients who tested positive for COVID-19 between January 4th, 2020 and May 3rd, 2022 were available in the study (Fig. 1). The study excluded patients for whom basic information was unavailable, such as comorbidities or demographic data (number of patients $n = 329300$), as well as those for whom erroneous values

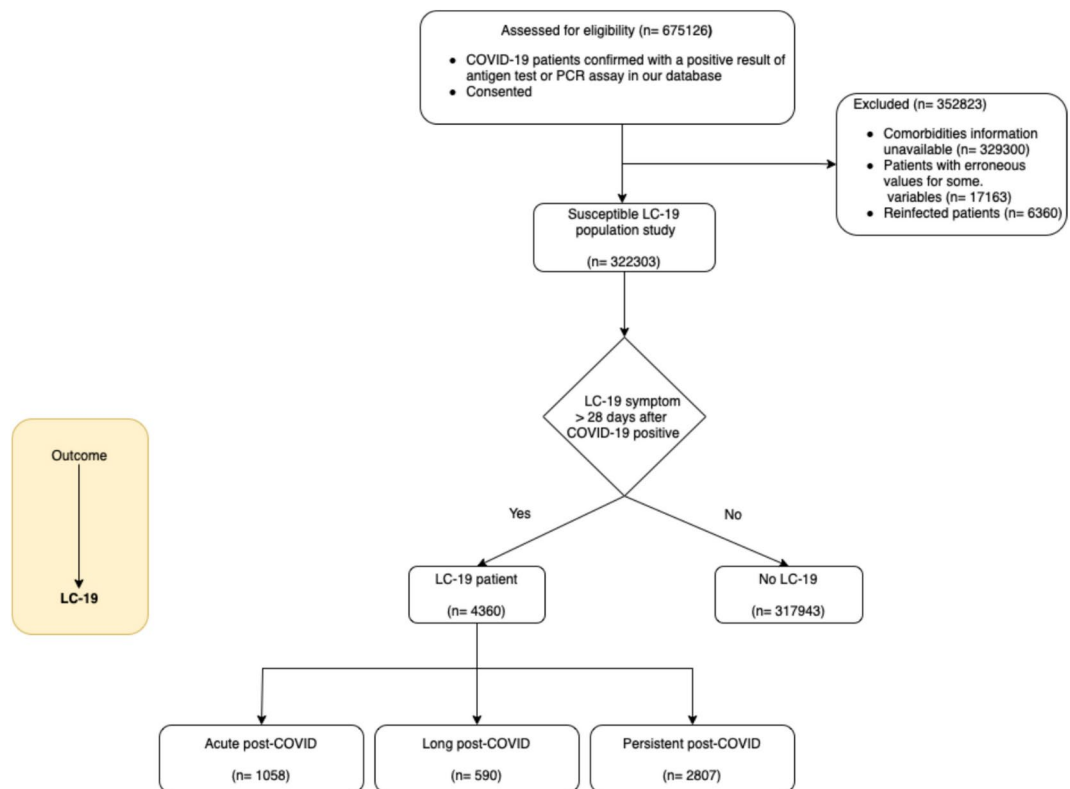


Fig. 1. Patient selection flow diagram. Flow diagram of the enrollment of subjects, disposition status, and how they are analysed in the study.

were present in any of the available records ($n = 17163$). Those erroneous values were identified by searching patients with at least one of the following constraints: negative age values, age values greater than 130, unknown sex, or multiple administrations of the vaccine dose in a single day. Furthermore, patients who have experienced reinfection are excluded from this study, as they are not the primary focus of interest ($n = 6360$). For all patients remaining ($n = 322303$), we checked if they exhibited LC-19 symptoms at least 28 days after the initial infection date to determine which patients had LC-19 ($n = 4360$). According to LC-19 categorisation definitions, a patient is assessed as an acute post-COVID case if any LC-19 symptom has persisted for more than 28 days but less than 12 weeks after the COVID-19 infection date ($n = 1058$). They are considered a long post-COVID case if they exhibit any LC-19 symptoms persistence between 12 and 24 weeks after the infection ($n = 590$), and a persistent post-COVID case if LC-19 symptoms persist beyond 24 weeks from the infection ($n = 2807$). The mean follow-up time for LC-19 patients is 358 days, whereas for those no LC-19 patients, the mean follow-up time is 160 days.

Since this study is a retrospective cohort analysis, the availability of patient data depended on the completeness of electronic health records, rather than on predefined data collection. Our inclusion criteria required patients to have a confirmed positive COVID-19 test, along with documented information on demographics, symptoms, comorbidities, and vaccination status. From the initial dataset of 675,126 patients with confirmed COVID-19 diagnoses, we identified that 329,300 patients lacked all data on comorbidities (16 variables) and demographic information (e.g., age and sex). Given that these variables were critical for our analyses, patients without this information did not meet the inclusion criteria and were therefore excluded. This exclusion was not a standard procedure for handling missing data within an established study population, but rather a necessary step to ensure that all included patients had sufficient data for valid statistical comparisons aligned with the study's objectives. Due to the large size of the dataset, the exclusion of these cases did not affect statistical power, but was essential to maintain the integrity and reliability of our analyses. The missing data pattern indicates that the lack of information was not missing at random or missing completely at random³³, but rather a consequence of structural gaps in the health records where the covariates did not exist. Including these cases would have introduced significant bias due to the absence of essential covariate data.

Data description and preprocessing

The demographic information and comorbidities of all patients in the population study were obtained from a database used for stratification purposes during COVID-19 pandemic. Available characteristics include age, sex, number of chronic diseases, number of affected systems, and various medical conditions such as diabetes mellitus, heart failure, chronic obstructive pulmonary disease, arterial hypertension, depression, ischaemic cardiomyopathy, stroke, renal insufficiency, cirrhosis, osteoporosis, osteoarthritis, arthritis, obesity, and asthma. A total of 322,303 patients have been studied; however, symptomatology records are available for

297,737 patients who have sought care in primary healthcare settings. This data includes features such as lack of appetite, nasal congestion, rhinorrhoea, low grade fever, eye symptoms, expectorate, sore throat, stomach pain, vomit, nasal discharge, chills and malaise. To perform the analyses that are based on the vaccination database, only those patients who received a vaccination 21 days prior to a confirmed diagnosis of COVID-19 infection were included in the study. Additionally, we eliminated patients whose schedule was inconclusive, for example patients that had received several doses on the same day. This left us with vaccine schedule information for 320,141 patients. Other data, such as medication (30,311 patients) or length of stay in each hospital department (15,829 patients), were not included in the analysis as they were deemed to be outside the scope of the study and would have significantly reduced the number of patients included.

Analysis of sociodemographic factors, comorbidities and symptoms between LC-19 patients and non-Long COVID-19 patients

Following the classification of LC-19 patients proposed by Fernández-de-las-Peñas et al.⁹ we conducted an exploratory analysis of sociodemographic factors, comorbidities and symptoms to seek for differences between non-Long COVID-19 and LC-19 patients (Table 1 and 2). The main difference between non-Long COVID-19 and LC-19 patients regarding sociodemographic factors was sex (Female %: 52.5% v. 61.1 % respectively). Being female increased the likelihood of developing LC-19 (AOR [95% CI] = 1.39 [1.31, 1.48], adjusted $p < 0.001$). No differences in sex composition were observed between the subtypes of LC-19 patients. The age median was very similar for non-Long COVID-19 and those with the condition (38 v. 37, respectively). However, we observed age differences among the subtypes of LC-19 patients; acute post-COVID patients were older than both long post-COVID and persistent post-COVID patients (adjusted $p < 0.001$).

Considering sex as an important factor in the development of LC-19 prompted us to analyse comorbidities and symptoms, disaggregated by sex, as illustrated in Fig. 2. In female patients, obesity (AOR [95% CI] = 1.28 [1.13, 1.45]) and depression (AOR [95% CI] = 1.38 [1.23, 1.54]) were identified as comorbidities associated with a statistically significant increased risk for LC-19. Conversely, dementia (AOR [95% CI] = 0.23 [0.1, 0.52]), arterial hypertension (AOR [95% CI] = 0.72 [0.62, 0.83]) and osteoporosis (AOR [95% CI] = 0.75 [0.61, 0.93]) were associated with a reduced risk of LC-19 (Fig. 2A). In male patients, no comorbidities were found to be associated with an increased risk of LC-19, while only stroke (AOR [95% CI] = 0.54 [0.29, 0.98]) was associated with a reduced risk of LC-19. We performed the same analysis disaggregated by LC-19 type of patient and no differences were observed (Supplementary Fig. 1). Additionally, crude ORs were obtained and are presented in the Supplementary Table 1.

Concerning differences in symptomatology, we noted a higher percentage of initial symptoms in LC-19 patients compared to non-Long COVID-19 patients (Table 2 and Supplementary Fig. 2). The five symptoms most significantly increasing the likelihood of LC-19 development were chest pain (AOR [95% CI] = 1.67 [1.41,

Characteristics	Non-LC-19	LC-19	Acute post-COVID	Long post-COVID	Persistent post-COVID
No. of individuals (N)	317943	4360	1058	590	2807
Age median (IQR)	38 (21, 52)	37 (23, 50)	48 (36, 58)	34 (19.25, 48)	33 (22, 45)
Sex					
Male (%)	151174 (47.5%)	1695 (38.9%)	413 (39%)	237 (40.2%)	1072 (38.2%)
Female (%)	166769 (52.5%)	2665 (61.1%)	645 (61%)	353 (59.8%)	1735 (61.8%)
Comorbidities					
Number of chronic diseases median (IQR)	2 (1, 4)	3 (1, 5)	4 (2, 6)	3 (1, 4)	2 (1, 4)
Number of systems affected median (IQR)	2 (1, 3)	2 (1, 4)	3 (2, 4)	2 (1, 3)	2 (1, 4)
Diabetes mellitus (%)	20585 (6.5%)	241 (5.5%)	112 (10.6%)	27 (4.6%)	107 (3.8%)
Heart failure (%)	3231 (1%)	26 (0.6%)	15 (1.4%)	4 (0.7%)	7 (0.2%)
COPD (%)	4869 (1.5%)	47 (1.1%)	19 (1.8%)	8 (1.4%)	21 (0.7%)
Arterial hypertension (%)	47036 (14.8%)	550 (12.6%)	233 (22%)	65 (11%)	263 (9.4%)
Depression (%)	29289 (9.2%)	524 (12%)	182 (17.2%)	74 (12.5%)	290 (10.3%)
ICM (%)	5823 (1.8%)	58 (1.3%)	24 (2.3%)	8 (1.4%)	27 (1%)
Stroke (%)	4019 (1.3%)	37 (0.8%)	12 (1.1%)	3 (0.5%)	23 (0.8%)
Renal insufficiency (%)	5950 (1.9%)	69 (1.6%)	24 (2.3%)	12 (2%)	34 (1.2%)
Cirrhosis (%)	6259 (2%)	93 (2.1%)	37 (3.5%)	8 (1.4%)	48 (1.7%)
Osteoporosis (%)	8765 (2.8%)	103 (2.4%)	46 (4.3%)	9 (1.5%)	49 (1.7%)
Osteoarthritis (%)	18534 (5.8%)	284 (6.5%)	106 (10%)	34 (5.8%)	151 (5.4%)
Arthritis (%)	4471 (1.4%)	73 (1.7%)	26 (2.5%)	8 (1.4%)	39 (1.4%)
Obesity (%)	28890 (9.1%)	461 (10.6%)	165 (15.6%)	40 (6.8%)	262 (9.3%)
Asthma (%)	30302 (9.5%)	466 (10.7%)	116 (11%)	70 (11.9%)	292 (10.4%)

Table 1. Demographic characteristics and comorbidities for COVID-19 patients (Non-LC-19) and different subtypes of long COVID-19 patients. Continuous data are reported as median with interquartile range (Q3-Q1), and categorical data are expressed as percentages (%). COPD is chronic obstructive pulmonary disease and ICM is Ischemic cardiomyopathy.

Characteristics	Non-LC-19	LC-19	Acute post-COVID	Long post-COVID	Persistent post-COVID
No. of individuals (N)	293,377	4360	1058	590	2807
Symptoms					
Chest pain (%)	12,117 (4.2%)	616 (14.1%)	266 (25.1%)	55 (9.3%)	309 (11%)
Muscle pain (%)	77,623 (27%)	1993 (45.7%)	611 (57.8%)	209 (35.4%)	1226 (43.7%)
Cough (%)	31,338 (10.9%)	1234 (28.3%)	494 (46.7%)	128 (21.7%)	650 (23.2%)
Lack of appetite (%)	9970 (3.5%)	518 (11.9%)	251 (23.7%)	49 (8.3%)	239 (8.5%)
Sleep problems (%)	7526 (2.6%)	471 (10.8%)	242 (22.9%)	43 (7.3%)	204 (7.3%)
Nasal congestion (%)	75,252 (26.2%)	1377 (31.6%)	352 (33.3%)	163 (27.6%)	896 (31.9%)
Dyspnea (%)	4630 (1.6%)	379 (8.7%)	210 (19.8%)	33 (5.6%)	147 (5.2%)
Rhinorrhoea (%)	87,434 (30.5%)	1553 (35.6%)	385 (36.4%)	205 (34.7%)	1008 (35.9%)
Low grade fever (%)	23,368 (8.1%)	411 (9.4%)	101 (9.5%)	57 (9.7%)	266 (9.5%)
Dizziness (%)	10,290 (3.6%)	523 (12%)	211 (19.9%)	58 (9.8%)	277 (9.9%)
Hyposmia (%)	41,460 (14.4%)	1427 (32.7%)	349 (33%)	127 (21.5%)	990 (35.3%)
Hypogeusia (%)	37,582 (13.1%)	1320 (30.3%)	338 (31.9%)	121 (20.5%)	901 (32.1%)
Headache (%)	87,205 (30.4%)	2158 (49.5%)	629 (59.5%)	250 (42.4%)	1337 (47.6%)
Tired (%)	16,097 (5.6%)	892 (20.5%)	402 (38%)	92 (15.6%)	432 (15.4%)
Eye symptoms (%)	1930 (0.7%)	102 (2.3%)	40 (3.8%)	15 (2.5%)	50 (1.8%)
Expectorate (%)	27,041 (9.4%)	776 (17.8%)	304 (28.7%)	89 (15.1%)	407 (14.5%)
Fever (%)	37,070 (12.9%)	706 (16.2%)	199 (18.8%)	98 (16.6%)	433 (15.4%)
Sore throat (%)	72,993 (25.4%)	1467 (33.6%)	409 (38.7%)	187 (31.7%)	918 (32.7%)
Stomach pain (%)	11,128 (3.9%)	503 (11.5%)	181 (17.1%)	54 (9.2%)	287 (10.2%)
Vomit (%)	7360 (2.6%)	284 (6.5%)	105 (9.9%)	29 (4.9%)	160 (5.7%)
Nasal discharge (%)	4558 (1.6%)	170 (3.9%)	72 (6.8%)	16 (2.7%)	92 (3.3%)
Chills (%)	23,910 (8.3%)	794 (18.2%)	290 (27.4%)	84 (14.2%)	450 (16%)
Rash (%)	2163 (0.8%)	124 (2.8%)	44 (4.2%)	12 (2%)	71 (2.5%)
Malaise (%)	859 (0.3%)	38 (0.9%)	20 (1.9%)	4 (0.7%)	16 (0.6%)

Table 2. Symptomatology for COVID-19 patients (Non-LC-19) and different subtypes of long COVID-19 patients. Categorical data are expressed as percentages (%).

1.99] for males and AOR [95% CI] = 1.58 [1.40, 1.79]) for females), hyposmia (AOR [95% CI] = 1.54 [1.27, 1.87] for males and AOR [95% CI] = 1.57 [1.36, 1.82]) for females), and stomach pain (AOR [95% CI] = 1.54 [1.28, 1.86] for males and AOR [95% CI] = 1.52 [1.34, 1.72]) for females), followed by rash (AOR [95% CI] = 1.46 [1, 2.12] for males and AOR [95% CI] = 1.73 [1.38, 2.18]) for females), and dyspnea (AOR [95% CI] = 1.72 [1.36, 2.17] for males and AOR [95% CI] = 1.35 [1.14, 1.59]) for females) (Fig. 2B). Regarding symptom sex differences, no significant variations were observed in the likelihood of experiencing a particular symptom during the primary COVID-19 infection and subsequently developing LC-19. We conducted the same analysis, stratified by LC-19 patient type, and found that for most symptoms, there were no significant differences among LC-19 patient subtypes (Supplementary Fig. 3). However, hyposmia showed a markedly stronger association with persistent LC-19 patients (AOR [95% CI] = 1.92 [1.66, 2.22], adjusted $p < 0.005$) than with acute post-COVID (AOR [95% CI] = 1.16 [0.93, 1.46], adjusted $p > 0.05$) or long post-COVID patients (AOR [95% CI] = 1.03 [0.72, 1.45], adjusted $p > 0.05$). Similarly, the association of chest pain was notably more pronounced in acute post-COVID patients (AOR [95% CI] = 2.08 [1.77, 2.46], adjusted $p < 0.001$) compared to those with persistent LC-19 (AOR [95% CI] = 1.42 [1.24, 1.62], adjusted $p < 0.001$) and long post-COVID (AOR [95% CI] = 1.28 [0.94, 1.75], adjusted $p > 0.05$). Crude ORs for symptoms were also obtained and are presented in the Supplementary Table 2.

Impact of COVID-19 hospitalisation on the development of LC-19

To understand the implications of COVID-19, the need of hospitalisation serve as a crucial indicator, shedding light on the disease's progression and severity across diverse patient profiles. This analysis focuses on the AORs for developing LC-19 among hospitalised and non-hospitalised patients, paying close attention to the interplay between sex and disease severity. We particularly differentiate among various LC-19 patient classifications through their impact on health outcomes. Illustrated in Fig. 3, our analysis offers a clear depiction of the odds associated with LC-19. Notably, hospitalised patients in the LC-19 category exhibit an AOR [95% CI] = 3.57 [3.22, 3.95], with an adjusted $p < 0.001$, adjusted by age and sex. Upon disaggregation by sex, an AOR [95% CI] = 4.49 [3.85, 5.23], with an adjusted $p < 0.001$, was observed for males and an AOR [95% CI] = 3.05 [2.66, 3.51], with an adjusted $p < 0.001$, was noted for females, suggesting a significantly higher likelihood of developing LC-19 overall compared to non-hospitalised counterparts. Specifically, acute post-COVID conditions present a markedly increased risk for hospitalised patients (AOR [95% CI] = 10.26 [8.83, 11.92], adjusted $p < 0.001$), especially for males (AOR [95% CI] = 15.56 [13.08, 20.97], adjusted $p < 0.001$) compared to females (AOR [95% CI] = 7.36 [6.02, 9.01], adjusted $p < 0.001$). Similarly, for persistent post-COVID conditions, a statistically significant higher

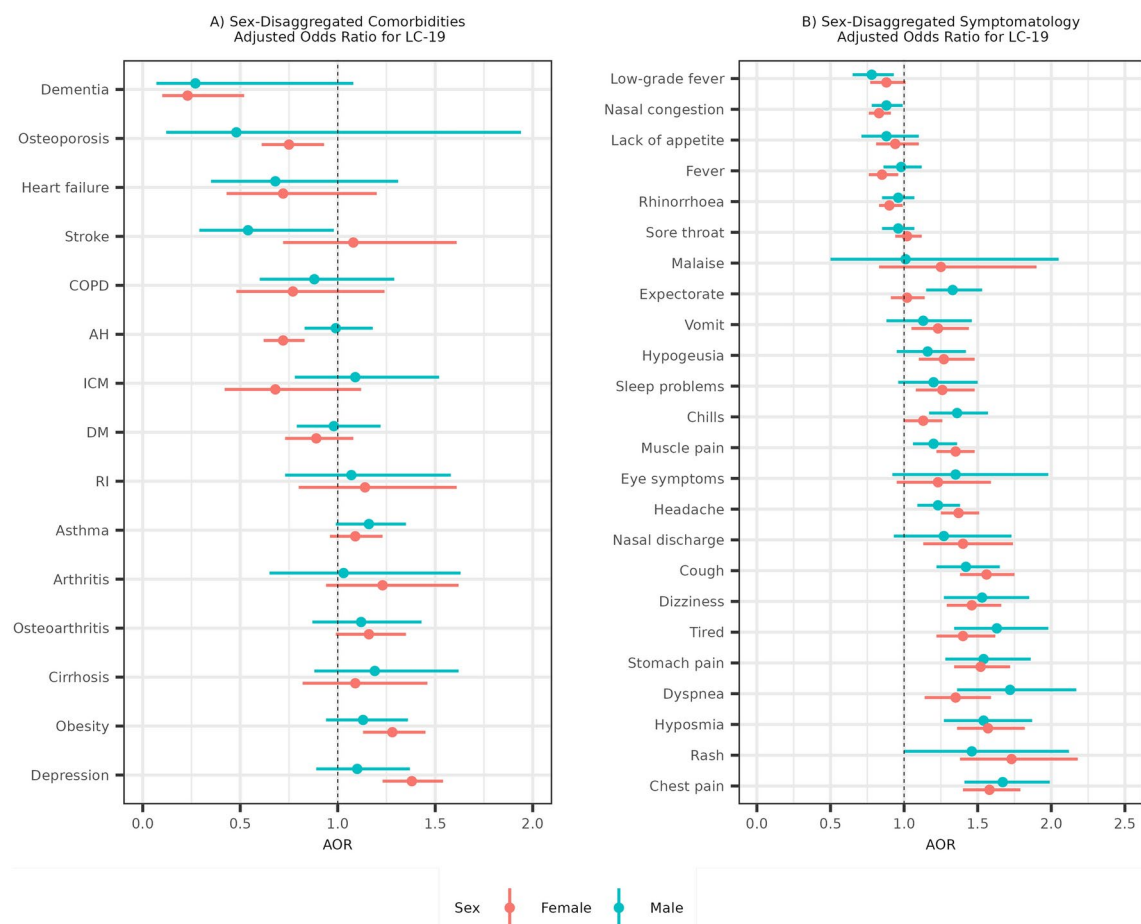


Fig. 2. (A) Adjusted odds ratios for developing LC-19 based on different comorbidities. (B) Adjusted odds ratios for developing LC-19 based on acute-phase COVID-19 symptoms. The female sex is coloured in red and the male sex in blue. Odds ratios are adjusted by age, the symptoms and the comorbidities. COPD is chronic obstructive pulmonary disease, AH is arterial hypertension, ICM is ischemic cardiomyopathy, DM is diabetes mellitus and RI is renal insufficiency.

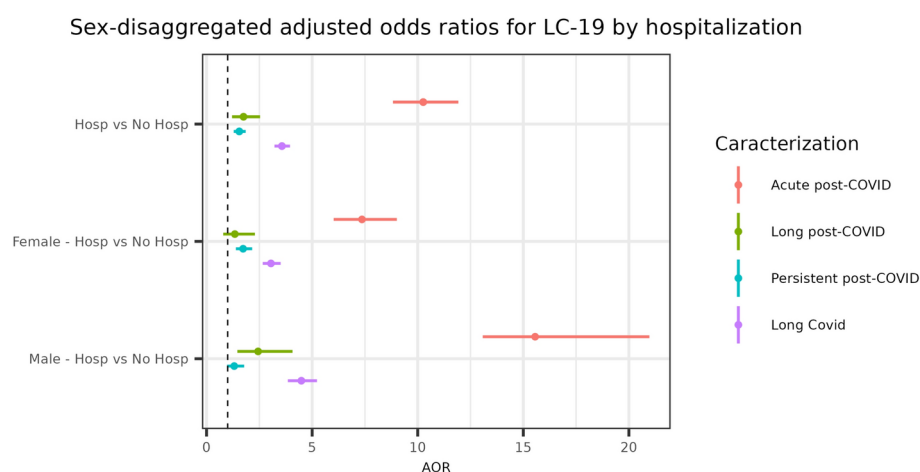


Fig. 3. Odds ratio for suffering each categorisation of LC-19 adjusted by age with and without sex-disaggregated.

risk is observed for hospitalised patients of both sex (AOR [95% CI] = 1.55 [1.29, 1.85], adjusted $p < 0.001$). For Long post-COVID conditions, hospitalised patients face a considerable risk (AOR [95% CI] = 1.75 [1.21, 2.53], adjusted $p < 0.005$), with this increased risk being significant solely for male patients (AOR [95% CI] = 2.44 [1.46, 4.07], adjusted $p < 0.001$). This suggests that hospitalisation is not only a high-risk factor for developing long post-COVID and persistent post-COVID manifestations of LC-19 classifications but is particularly crucial for acute post-COVID patients. The AOR for acute post-COVID patients is significantly higher than for other LC-19 classifications. In order to demonstrate the inherent risk associated with hospitalisation, crude ORs are presented in the Supplementary Table 3.

Effects of vaccination on LC-19

We examined the effects of COVID-19 vaccination, including the impact of dosage, patterns, and brand differences on the course of LC-19. Supplementary Fig. 4 shows the vaccine administration flow in the cohort under study. Only 15.95% of the cohort is unvaccinated. The most commonly used vaccine brands, in our data, for the first administration are Pfizer (72.39%) and Moderna (15%), followed by Astra Zeneca (8.22%) and Janssen (4.39%). The initial dominance of Pfizer decreases over time. This is shown by the transition from first to second doses. Pfizer still leads, but with a reduced majority of 67.07% of the second dose administration, and it is followed by Moderna (24.25%), Astra Zeneca (8.68%) and Janssen ($< 0.01\%$). Meanwhile, 33.6% of the cohort did not receive a second dose. Furthermore, 80.46% of the population has not yet received a third dose. Among the available brands, Moderna has administered the highest proportion of third doses at 80.20%, followed by Pfizer at 19.79% and Astra Zeneca (0.01%). No Janssen doses were recorded during the administration of the third dose.

When evaluating the protective effect of vaccination in relation to the health impact of COVID-19, the results are significantly influenced by vaccination status. Table 3 demonstrates a clear protective effect of vaccination vs no vaccine against LC-19, after adjusting by age and sex the AOR [95% CI] = 0.10 [0.09, 0.12], and adjusted $p < 0.001$. The analysis of vaccine doses and their impact on LC-19 incidence reveals a notable more pronounced protective effect when comparing two doses against one dose (AOR [95% CI] = 0.38 [0.29, 0.50], adjusted $p < 0.001$) and even three doses against both two doses (AOR [95% CI] = 0.39 [0.27, 0.58], adjusted $p < 0.001$) and one dose (AOR [95% CI] = 0.11 [0.06, 0.18], adjusted $p < 0.001$). We also observed that mRNA vaccines provided a more protective effect against LC-19 than viral vector-based vaccines (AOR [95% CI] = 0.48 [0.37, 0.63], adjusted $p < 0.001$). We performed the analysis of mRNA vaccines vs viral vector-based vaccines disaggregated by sex and found this effect to be statistically significant only in women (Supplementary Fig. 5). Considering the effect of mRNA vaccines, we studied a combination of mRNA and viral vector-based vaccines, comparing them against two doses of either mRNA or viral vector-based vaccines alone. A mix of mRNA and viral vector-based vaccines in any order was more protective than two doses of viral vector-based vaccines (AOR [95% CI] = 0.3 [0.15, 0.58], adjusted $p < 0.001$) and as protective as two doses of mRNA vaccines against LC-19 (AOR [95% CI] = 0.77 [0.41, 1.46], adjusted $p = 0.43$). Additionally, in order to observe the inherent risk associated with each condition mentioned before, crude ORs are available in the Supplementary Table 4.

Furthermore, we investigated the effects of various COVID-19 vaccination schedules, segmented by vaccine brand, on LC-19 incidence (Fig. 4). AORs were utilised to determine the relative protection offered by specific vaccination schedules. Figure 4 summarises the protective effects of different vaccination schedules against LC-

	LC-19	No LC-19	AOR [95% CI]	Adj p
	N (%)			
No vaccinated (Ref)	3916 (2.4%)	160,007 (97.6%)	0.10 [0.09, 0.12]	< 0.001
Vaccinated	413 (0.3%)	155,805 (99.7%)		
COVID-19 viral-based vaccine (Ref)	87 (0.71%)	12,096 (99.29%)	0.48 [0.37, 0.63]	< 0.001
mRNA-based vaccine	324 (0.23%)	140,671 (99.77%)		
2 COVID viral-based vaccine doses (Ref)	62 (0.70%)	8799 (99.30%)	0.30 [0.15, 0.58]	< 0.001
COVID viral-based dose + mRNA-based dose	10 (0.20%)	5073 (99.80%)		
2 mRNA-based vaccine doses (Ref)	242 (0.23%)	105,482 (99.77%)	0.77 [0.41, 1.46]	0.43
COVID viral-based dose + mRNA-based dose	10 (0.20%)	5073 (99.80%)		
No. of doses				
0 (Ref) vs	3916 (2.4%)	160,007 (97.6%)	–	
1	66 (0.6%)	11,251 (99.4%)	0.25 [0.19, 0.32]	< 0.001
2	314 (0.3%)	119,354 (99.7%)	0.11 [0.09, 0.12]	< 0.001
3	33 (0.1%)	25,200 (99.9%)	0.05 [0.03, 0.07]	< 0.001
1 (Ref) vs	66 (0.6%)	11,251 (99.4%)	–	
2	314 (0.3%)	119,354 (99.7%)	0.38 [0.29, 0.50]	< 0.001
3	33 (0.1%)	25,200 (99.9%)	0.11 [0.06, 0.18]	< 0.001
2 (Ref) vs	314 (0.3%)	119,354 (99.7%)	–	
3	33 (0.1%)	25,200 (99.9%)	0.39 [0.27, 0.58]	< 0.001

Table 3. Table of vaccination differences for non LC-19 and LC-19 patients.

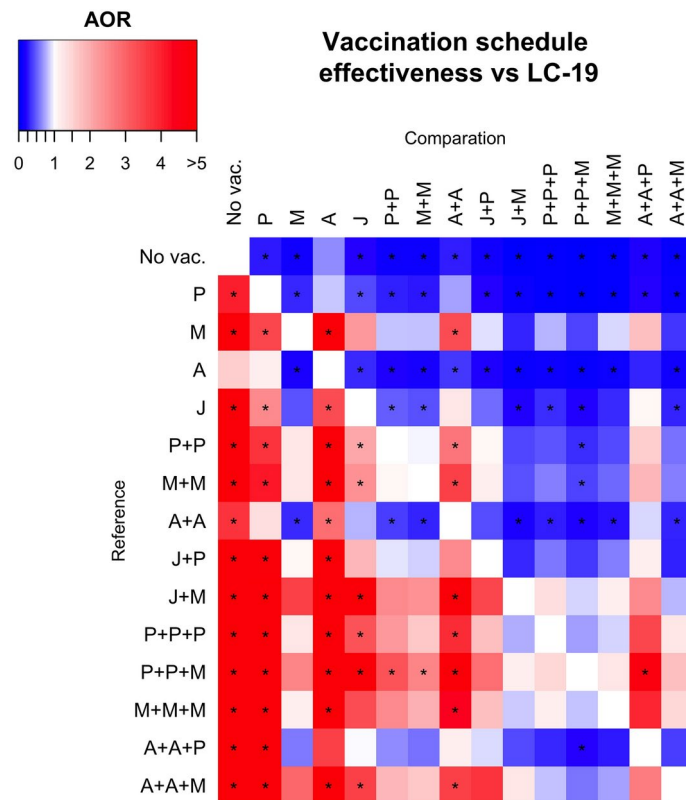


Fig. 4. Adjusted odds ratios (AORs) for various vaccination schedules against LC-19. Vaccination schedules are represented by combinations of letters that reflect the order in which the vaccines were administered to participants. With A for AstraZeneca, P for Pfizer, M for Moderna, and J for Janssen vaccines. Blue squares represent greater protective behavior indicated by lower adjusted odds ratios (AOR), when comparing the reference schedules (rows) with the comparison ones (columns), being then the comparison one more protective. In contrast, red squares signify a higher AOR, indicating less protective effect against the outcome. If the adjusted p-value for the AOR is less than 0.05, the comparison is flagged with a star (*).

19. Initially, it was found that any vaccination schedule provides more protection against the development of LC-19 compared to being unvaccinated. The three most protective schedules were Pfizer + Pfizer + Moderna, Pfizer + Pfizer + Pfizer and Janssen + Moderna. Conversely, the least protective vaccine schedules against LC-19 were a single dose of AstraZeneca (no statistically significant protection against LC-19 comparing with no vaccine administration), a single dose of Pfizer, and two doses of AstraZeneca. Exact values of AORs for each analysis is presented in Supplementary Table 5 and, in order to observe the inherent risk of each vaccine schedule comparison, all crude ORs are also available in Supplementary Table 6.

LC-19 calculator tool

A predictive model was developed to predict the probability of suffering from LC-19, depending on predictors like comorbidities, sex, age, the vaccination schedule up to 21 days before the positive diagnosis of COVID-19, and the symptomatology of the individual at the initial diagnosis of the COVID-19 infection and up to 28 days after the positive test for COVID-19. Only data collected prior to the LC-19 diagnosis were considered for training the model. Only patients with all the above information available were considered for this objective. The training dataset (199527 no LC-19 patients and 3031 LC-19 patients) was used to train the model and the test dataset (85511 no LC-19 patients and 1298 LC-19 patients) was used to evaluate the model.

IPIP and Smote were used to deal with the imbalanced proportion of LC-19 and no LC-19 patients together with logistic regressions as ML algorithms. IPIP with logistic regression obtained a balanced accuracy of 73% and ROC-AUC of 0.80 (Supplementary Fig. 6) on the test dataset and for Smote with logistic regression we obtained a balanced accuracy of 65% and a ROC-AUC of 0.79. Then, we used IPIP with logistic regression models to create an online LC-19 calculator that is accessible at ([LC-19 Calculator](#)).

In order to utilise the calculator to get a probability of suffering LC-19, it is necessary to input the patient's sex, age, comorbidities, and symptomatology during the acute-phase of COVID-19, as well as their vaccination schedule prior to 21 days of infection.

Discussion

In this study, we investigated the demographic factors, comorbidities, and symptoms of the acute phase of COVID-19 as risk factors for LC-19. Our study offers valuable insights into the factors that influence the

development of long COVID. Notably, our data support that female sex is associated with a higher likelihood of developing long COVID. Another risk factor is the number of symptoms experienced during the acute phase of COVID-19; patients who exhibit a high number of symptoms are more likely to develop long COVID. Additionally, the need for hospitalisation during the acute phase significantly increases the risk of developing long COVID, emphasising the severity of the initial infection as a key factor in long-term outcomes. Our analysis also shows no strongly significant differences in symptoms and comorbidities across different subtypes of LC-19 patients, suggesting that the manifestation of long COVID is consistent among LC-19 subtypes of patients. Crucially, vaccination before acute SARS-CoV-2 infection significantly reduces the likelihood of developing long COVID. Among the vaccines, mRNA vaccines provide greater protection against long COVID compared to viral vector-based vaccines. Furthermore, a combination of mRNA and viral vector-based vaccines, in any sequence, offers more protection than two doses of viral vector-based vaccines and is as effective as two doses of mRNA vaccines.

Despite the valuable insights provided by our study, several limitations must be acknowledged. First, the number of individuals with LC-19 in our sample is relatively small compared to non-LC-19 patients. However, the majority of our results are consistent with existing studies. Second, the diagnosis of LC-19 in our study was not clinically confirmed. Instead, we relied on a database derived from health registry data from COVID-19 patients, identifying LC-19 cases based on reported symptoms and follow-up consultations with clinicians. This method may introduce bias due to variability in symptom interpretation, as well as differences in patients' willingness to report symptoms or seek medical care. Another limitation is the lack of detailed historical data on comorbidities. While the database included pre-existing conditions, we lacked information on the duration and timing of these conditions. For example, the database registered whether a patient had experienced a heart attack prior to infection but did not capture when it occurred. Another limitation of our study did not account for the specific variants of the SARS-CoV-2 virus. Variants may have different impacts on both the likelihood and severity of Long COVID, which was not addressed in our analysis^{34–36}.

According to our analysis of sociodemographic factors, in the cohort under study female sex is associated with a higher likelihood of developing long COVID. This finding is consistent with other published studies that have identified female sex as a significant risk factor for LC-19^{14,25–29}. Moreover, the impact of female sex appears to be uniform across all subtypes of LC-19 patients, as illustrated in Supplementary Fig. 1. Regarding the risk factors associated with LC-19, we found that comorbidities such as depression and obesity increase the risk of LC-19 in females. This finding aligns with a study predominantly involving female participants (96%) that also identified depression before COVID as a significant risk factor for developing long COVID³⁷. Other studies also pointed out obesity as a risk factor for developing LC-19^{12,38–40}. Additionally, another study in Denmark concluded that obesity and depression are risk factors for LC-19⁴¹. On the other hand, we found that female patients with dementia have a lower likelihood of developing LC-19. It is well-documented that patients with dementia have higher odds of dying from COVID-19^{24,42,43}. Additionally, COVID-19 appears to increase the likelihood of developing dementia after infection, suggesting a potential association between COVID-19 and dementia^{44–46}. However, to the best of our knowledge, there is no evidence indicating whether patients with dementia have higher or lower odds of developing LC-19. We hypothesise that our results could be influenced by the sample size of patients with dementia and the fact that underreporting and undiagnosing are more common among patients with dementia^{47–49}.

Regarding acute-phase symptoms, we found that patients exhibiting several symptoms during the acute phase are at an increased likelihood of developing LC-19 after a COVID-19 infection. This finding aligns with a multicenter study that identified a greater number of symptoms at hospital admission as the most significant risk factor for developing more symptoms after COVID-19²⁵. In detail, we found that the acute-phase symptoms most strongly associated with an increased likelihood of developing LC-19 are chest pain, rash, hyposmia, dyspnea, and stomach pain. Our findings align with a Polish study, indicating that chest pain and dyspnea are risk factors for the development of LC-19²⁹. The same symptoms that we previously identified were also reported as risk factors in a large cohort of non-hospitalised COVID-19 patients¹². Some studies which examine cohorts of individuals from different age groups have identified chest pain, hyposmia and dyspnea as risk factors for suffering from LC-19, not only in adults but also in children^{50,51}.

Another relevant question, due to ambiguous conclusions, is whether the severity of the COVID-19 infection influences the development of LC-19. Some studies have concluded that the severity of the acute disease does not significantly contribute to the risk of developing long COVID^{14,52,53}. However, the majority of studies have found that the severity or need for hospitalisation during the acute phase is associated with a higher likelihood of developing LC-19 symptoms^{54–58}. Our results align with these latter studies, suggesting a significantly higher likelihood of developing LC-19 for hospitalised patients compared to non-hospitalised counterparts. This effect is particularly strong in the acute-post COVID phase and moderates over time, as observed for persistent post-COVID in Fig. 3.

Several publications have investigated the impact of vaccination schedules prior to acute SARS-CoV-2 infection in LC-19^{17,19,59–64}. These studies support the notion of a protective effect of COVID-19 vaccination against long COVID. Our findings align with these conclusions, indicating that vaccinated individuals have a lower risk of developing LC-19 compared to unvaccinated individuals (AOR [95% CI] = 0.10 [0.09, 0.12]). Furthermore, consistent with some of these previous studies^{59,64}, our research suggests that receiving two vaccine doses provides greater protection against LC-19 compared to receiving only one dose (AOR [95% CI] = 0.38 [0.29, 0.58]). In our investigation, we observed that three doses exhibited a more pronounced protective effect against LC-19 than both two doses (AOR [95% CI] = 0.39 [0.27, 0.58]) and one dose (AOR [95% CI] = 0.11 [0.06, 0.18]), respectively. Therefore, the third dose could be beneficial for preventing LC-19.

Another question arises as to whether mRNA vaccines and viral vector-based vaccines exhibit differences in the likelihood of developing LC-19. Some studies^{65,66} provide evidence suggesting a more pronounced

protective effect of mRNA vaccines compared to viral vector-based ones in terms of immune response, efficacy, and effectiveness. However, there is limited investigation on this question regarding LC-19. To the best of our knowledge, we found only one publication indicating a slightly stronger preventive effect for BNT162b2 (Pfizer/mRNA) compared to ChAdOx1 (AstraZeneca/viral vector-based)¹⁹. Consistently, our findings in the cohort corroborate and expand this observation; while both types of vaccines confer protection against LC-19, mRNA vaccines offer a higher degree of protection than viral vector-based vaccines (AOR [95% CI] = 0.48 [0.37, 0.63]). However, according to our results, this effect is only statistically significant in women (Supplementary Fig. 5). Additionally, we concluded that, regarding two doses, a mix of mRNA and viral vector-based vaccines in any order was more protective than two doses of viral vector-based vaccines and as protective as two doses of mRNA vaccines against LC-19. These findings underscore the critical role of vaccination, especially mRNA vaccines, in preventing long COVID and support the use of mixed vaccine regimens in protection. Our study highlights the necessity for ongoing public health efforts to promote vaccination and manage acute COVID-19 symptoms to lessen the burden of long COVID. We also presented a detailed analysis of the effects of various vaccination schedules on LC-19 in Fig. 4. Compared to not being vaccinated, receiving any vaccination schedule, except a single dose of AstraZeneca, significantly reduces the risk of developing LC-19. Among individuals vaccinated with one dose, those who received Moderna had a lower probability of developing LC-19 than the others. For those with two doses, the Pfizer+Pfizer, Moderna+Moderna, and Janssen+Moderna schedules were the best ones in terms of reducing risk. In individuals with three doses, the Pfizer+Pfizer+Moderna schedule showed the lowest probability of developing LC-19. As illustrated in the figure and previous analysis, individuals with more doses, particularly mRNA vaccines, demonstrated a reduced risk of developing LC-19. These findings are consistent with other studies and provide additional evidence supporting the most protective vaccination combinations^{19,59,64–66}. However, regarding the three doses effects on LC-19 there are fewer studies^{67,68}. Our results align with those studies on the most protective effect of three doses on LC-19 development and detail the schedules and brand combinations.

In this study, we developed a model to predict the likelihood of developing LC-19 during the acute phase of COVID-19 by considering demographic data, pre-existing comorbidities, symptoms at the initial diagnosis, and vaccination record prior to 21 days of infections. We have made this tool available at ([LC-19 Calculator](#)), and it achieved a balanced accuracy of 73% and an ROC-AUC of 0.80. Other studies have also developed predictive models using machine learning algorithms and obtained similar results^{69,70}. For instance, the model developed by Antony et al. achieved an ROC-AUC of 0.76, although they did not include vaccination records⁶⁹. Kessler et al. developed a more complex predictive model using light gradient boosting machine algorithm, which achieved an ROC-AUC of 0.84, although they included information about SARS-CoV-2 variants⁷⁰. We aimed to develop a simple and practical model using information collected at the time of COVID-19 diagnosis, allowing clinicians to use it as an effective triage tool. Given the model's accuracy and the study's limitations, it is clear that clinical decisions should not rely solely on this tool. However, we believe it can contribute to closer monitoring of higher-risk individuals, especially in the Murcia Region.

The external validity of our study is supported by the alignment of many findings with previous research. However, several factors could limit its generalisability. Our study is based on a retrospective cohort analysis using health registry data from a specific period. While the large sample size strengthens our conclusions compared to smaller cohort studies, the quality and granularity of our data are not directly comparable to prospective follow-up studies. This study covers the period from January 4, 2020, to May 3, 2022, and the patterns observed may have evolved due to the emergence of new SARS-CoV-2 variants and changes in vaccination strategies. Additionally, our sample is drawn from a single region, which may limit the external validity of our findings. Differences in demographics, healthcare systems, and pandemic response measures between regions may affect the applicability of our results to other populations. To enhance generalisability, future research should include larger, more diverse populations and examine the impact of SARS-CoV-2 variants and vaccines on Long COVID risk factors across broader contexts.

Data availability

All the data used in this study were retrieved from the Servicio Murciano de Salud (SMS). All data produced in the present study are available upon reasonable request to the corresponding authors (alejandrocisterna@um.es) and the approval by SMS. We developed a web to predict LC-19 probabilities <https://provia.inf.um.es/longcovid>.

Received: 18 November 2024; Accepted: 17 March 2025

Published online: 28 March 2025

References

- Hu, B., Guo, H., Zhou, P. & Shi, Z.-L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **19**(3), 141–154. <https://doi.org/10.1038/s41579-020-00459-7> (2021).
- WHO Coronavirus (COVID-19) Dashboard. <https://covid19.who.int>.
- Bergman, J., Ballin, M., Nordström, A. & Nordström, P. Risk factors for COVID-19 diagnosis, hospitalization, and subsequent all-cause mortality in Sweden: A nationwide study. *Eur. J. Epidemiol.* **36**(3), 287–298. <https://doi.org/10.1007/s10654-021-00732-w> (2021).
- Dessie, Z. G. & Zewotir, T. Mortality-related risk factors of COVID-19: A systematic review and meta-analysis of 42 studies and 423,117 patients. *BMC Infect. Dis.* **21**(1), 855. <https://doi.org/10.1186/s12879-021-06536-3> (2021).
- Nasserie, T., Hittle, M. & Goodman, S. N. Assessment of the frequency and variety of persistent symptoms among patients with COVID-19: A systematic review. *JAMA Netw. Open* **4**(5), 2111417. <https://doi.org/10.1001/jamanetworkopen.2021.11417> (2021).
- Bowyer, R. C. E. et al. The CONVALESCENCE Study: Characterising patterns of COVID-19 and long COVID symptoms: Evidence from nine UK longitudinal studies. *Eur. J. Epidemiol.* **38**(2), 199–210. <https://doi.org/10.1007/s10654-022-00962-6> (2023).

7. Vivaldi, G. et al. Long-term symptom profiles after COVID-19 vs other acute respiratory infections: an analysis of data from the COVidence UK study. *eClinicalMedicine*. **65**, 102251. <https://doi.org/10.1016/j.eclinm.2023.102251> (2023).
8. Davis, H. E., McCorkell, L., Vogel, J. M. & Topol, E. J. Author correction: Long COVID: major findings, mechanisms and recommendations. *Nat. Rev. Microbiol.* **21**(6), 408–408. <https://doi.org/10.1038/s41579-023-00896-0> (2023).
9. Fernández-de-las-Peñas, C., Palacios-Ceña, D., Gómez-Mayordomo, V., Cuadrado, M. L. & Florencio, L. L. Defining post-COVID symptoms (post-acute COVID, long COVID, persistent post-COVID): An integrative classification. *Int. J. Environ. Res. Public Health* **18**(5), 2621. <https://doi.org/10.3390/ijerph18052621> (2021).
10. Altmann, D. M., Whettlock, E. M., Liu, S., Arachchillage, D. J. & Boyton, R. J. The immunology of long COVID. *Nat. Rev. Immunol.* **23**(10), 618–634. <https://doi.org/10.1038/s41577-023-00904-7> (2023).
11. Byambasuren, O., Stehlik, P., Clark, J., Alcorn, K. & Glasziou, P. Effect of covid-19 vaccination on long covid: Systematic review. *BMJ Med.* **2**(1), 000385. <https://doi.org/10.1136/bmjmed-2022-000385> (2023).
12. Subramanian, A. et al. Symptoms and risk factors for long COVID in non-hospitalized adults. *Nat. Med.* **28**(8), 1706–1714. <https://doi.org/10.1038/s41591-022-01909-w> (2022).
13. Sieurin, J., Brandén, G., Magnusson, C., Hergens, M.-P. & Kosidou, K. A population-based cohort study of sex and risk of severe outcomes in covid-19. *Eur. J. Epidemiol.* **37**(11), 1159–1169. <https://doi.org/10.1007/s10654-022-00919-9> (2022).
14. ...Bai, F. et al. Female gender is associated with long COVID syndrome: A prospective cohort study. *Clin. Microbiol. Infect.* **28**(4), 611–e9. <https://doi.org/10.1016/j.cmi.2021.11.002> (2022).
15. Sylvester, S. V. et al. Sex differences in sequelae from COVID-19 infection and in long COVID syndrome: a review. *Curr. Med. Res. Opin.* **38**(8), 1391–1399. <https://doi.org/10.1080/03007995.2022.2081454> (2022).
16. Mohr, N. M. et al. Presence of symptoms 6 weeks after COVID-19 among vaccinated and unvaccinated US healthcare personnel: A prospective cohort study. *BMJ Open* **13**(2), 063141. <https://doi.org/10.1136/bmjopen-2022-063141> (2023).
17. Notarte, K. I. et al. Impact of COVID-19 vaccination on the risk of developing long-COVID and on existing long-COVID symptoms: A systematic review. *eClinicalMedicine*. **53**, 101624. <https://doi.org/10.1016/j.eclinm.2022.101624> (2022).
18. Ayoubkhani, D. et al. Trajectory of long covid symptoms after covid-19 vaccination: Community based cohort study. *BMJ*. <https://doi.org/10.1136/bmj-2021-069676> (2022).
19. Català, M. et al. The effectiveness of COVID-19 vaccines to prevent long COVID symptoms: Staggered cohort study of data from the UK, Spain, and Estonia. *Lancet Respir. Med.* [https://doi.org/10.1016/S2213-2600\(23\)00414-9](https://doi.org/10.1016/S2213-2600(23)00414-9) (2024).
20. Sánchez-de Prada, L. et al. Impact on the time elapsed since SARS-CoV-2 infection, vaccination history, and number of doses, on protection against reinfection. *Sci. Rep.* **14**(1), 353. <https://doi.org/10.1038/s41598-023-50335-6> (2024).
21. Nielsen, K. F. et al. Vaccine effectiveness against SARS-CoV-2 reinfection during periods of alpha, delta, or omicron dominance: A danish nationwide study. *PLoS Med.* **19**(11), 1004037. <https://doi.org/10.1371/journal.pmed.1004037> (2022).
22. Von Elm, E. et al. The strengthening the reporting of observational studies in epidemiology (strobe) statement: Guidelines for reporting observational studies. *Lancet* **370**(9596), 1453–1457 (2007).
23. Disease Control, C., Prevention: Post-COVID Conditions. <https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects/index.html> (Centers for Disease Control and Prevention, 2023).
24. Cisterna-García, A. et al. A predictive model for hospitalization and survival to covid-19 in a retrospective population-based study. *Sci. Rep.* **12**(1), 18126 (2022).
25. Fernández-de-las-Peñas, C. et al. Female sex is a risk factor associated with long-term post-COVID related-symptoms but not with COVID-19 symptoms: The LONG-COVID-EXP-CM multicenter study. *J. Clin. Med.* **11**(2), 413. <https://doi.org/10.3390/jcm11020413> (2022).
26. Rodríguez Onieva, A., Soto Castro, C. A., García Morales, V., Aneri Vacas, M. & Hidalgo Requena, A. Long COVID: Factors influencing persistent symptoms and the impact of gender. *Medicina de Familia. SEMERGEN*. **50**(5), 102208. <https://doi.org/10.1016/j.semerg.2024.102208> (2024).
27. Cohen, J. & Van Der Meulen Rodgers, Y. An intersectional analysis of long COVID prevalence. *Int. J. Equity Health*. **22**(1), 261. <https://doi.org/10.1186/s12939-023-02072-5> (2023).
28. Asadi-Pooya, A. A. et al. Risk factors associated with long COVID syndrome: A retrospective study. *Iran. J. Med. Sci.* <https://doi.org/10.30476/ijms.2021.92080.2326> (2021).
29. Chudzik, M. et al. Long-covid clinical features and risk factors: A retrospective analysis of patients from the stop-covid registry of the polocov study. *Viruses* **14**(8), 1755 (2022).
30. Chongsuvivatwong, V. epiDisplay: Epidemiological Data Display Package. R package version 3.5.0.2. <https://CRAN.R-project.org/package=epiDisplay> (2022).
31. Kuhn, M. Building predictive models in r using the caret package. *J. Stat. Softw.* **28**, 1–26. <https://doi.org/10.18637/jss.v028.i05> (2008).
32. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. Smote: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).
33. Kang, H. The prevention and handling of the missing data. *Korean J. Anesthesiol.* **64**(5), 402–406. <https://doi.org/10.4097/kjae.2013.64.5.402> (2013).
34. Fano-Sizgorich, D. et al. Risk of death, hospitalization and intensive care unit admission by SARS-CoV-2 variants in Peru: A retrospective study. *Int. J. Infect. Dis.* **127**, 144–149. <https://doi.org/10.1016/j.ijid.2022.12.020> (2023).
35. ...Hedberg, P. et al. In-hospital mortality during the wild-type, alpha, delta, and omicron SARS-CoV-2 waves: A multinational cohort study in the EuCARE project. *Lancet Reg. Health-Europe* **38**, 100855. <https://doi.org/10.1016/j.lanepe.2024.100855> (2024).
36. Hughes, T. D. et al. The effect of SARS-CoV-2 variant on respiratory features and mortality. *Sci. Rep.* **13**(1), 4503. <https://doi.org/10.1038/s41598-023-31761-y> (2023).
37. Wang, S. et al. Associations of depression, anxiety, worry, perceived stress, and loneliness prior to infection with risk of post-COVID-19 conditions. *JAMA Psychiat.* **79**(11), 1081. <https://doi.org/10.1001/jamapsychiatry.2022.2640> (2022).
38. Vimercati, L. et al. Association between long COVID and overweight/obesity. *J. Clin. Med.* **10**(18), 4143. <https://doi.org/10.3390/jcm10184143> (2021).
39. Debski, M. et al. Post-COVID-19 syndrome risk factors and further use of health services in east England. *PLoS Glob. Public Health*. **2**(11), 0001188. <https://doi.org/10.1371/journal.pgph.0001188> (2022).
40. Tenforde, M. W. et al. Symptom duration and risk factors for delayed return to usual health among outpatients with COVID-19 in a multistate health care systems network - United States, March-June 2020. *MMWR Morbid. Mortal. Wkly. Rep.* **69**(30), 993–998. <https://doi.org/10.15585/mmwr.mm6930e1> (2020).
41. Jakobsen, K. D., O'Regan, E., Svalgaard, I. B. & Hviid, A. Machine learning identifies risk factors associated with long-term sick leave following COVID-19 in Danish population. *Commun. Med.* **3**(1), 188. <https://doi.org/10.1038/s43856-023-00423-5> (2023).
42. Zhang, Q. et al. COVID-19 case fatality and Alzheimer's disease. *J. Alzheimers Dis.* **84**(4), 1447–1452. <https://doi.org/10.3233/JAD-215161> (2021).
43. Tahira, A. C., Verjovski-Almeida, S. & Ferreira, S. T. Dementia is an age-independent risk factor for severity and death in COVID-19 inpatients. *Alzheimer's Dementia*. **17**(11), 1818–1831. <https://doi.org/10.1002/alz.12352> (2021).
44. Magusali, N. et al. A genetic link between risk for Alzheimer's disease and severe COVID-19 outcomes via the OAS1 gene. *Brain* **144**(12), 3727–3741. <https://doi.org/10.1093/brain/awab337> (2021).

45. The Management Group of the EAN Dementia and Cognitive Disorders ScientificPanel. et al. Dementia and COVID-19, a bidirectional liaison: Risk factors, biomarkers, and optimal health care. *J. Alzheimer's Dis.* **82**(3), 883–898. <https://doi.org/10.3233/JAD-210335> (2021).
46. Shan, D., Wang, C., Crawford, T. & Holland, C. Temporal Association between COVID-19 Infection and Subsequent New-Onset Dementia in Older Adults: A Systematic Review and Meta-Analysis. <https://doi.org/10.2139/ssrn.4716751>. <https://www.ssrn.com/abstract=4716751> Accessed 23 May 2024. (2024).
47. Hodgson, N. A., Gitlin, L. N., Winter, L. & Czekanski, K. Undiagnosed illness and neuropsychiatric behaviors in community residing older adults with dementia. *Alzheimer Dis. Assoc. Disord.* **25**(2), 109–115. <https://doi.org/10.1097/WAD.0b013e3181f8520a> (2011).
48. Löppönen, M. K. et al. Undiagnosed diseases in patients with dementia—a potential target group for intervention. *Dement. Geriatr. Cogn. Disord.* **18**(3), 321–329. <https://doi.org/10.1159/000080126> (2004).
49. Thorpe, C. T. et al. Receipt of monitoring of diabetes mellitus in older adults with comorbid dementia. *J. Am. Geriatr. Soc.* **60**(4), 644–651. <https://doi.org/10.1111/j.1532-5415.2012.03907.x> (2012).
50. Sudre, C. H. et al. Attributes and predictors of long covid. *Nat. Med.* **27**(4), 626–631 (2021).
51. Morello, R., Mariani, F., Mastrantonio, L., De Rose, C., Zampino, G., Munblit, D., Sigfrid, L., Valentini, P. & Buonsenso, D. Risk factors for post-covid-19 condition (long covid) in children: a prospective cohort study. *EClinicalMedicine* **59** (2023).
52. Townsend, L. et al. Persistent fatigue following SARS-CoV-2 infection is common and independent of severity of initial infection. *PLoS ONE*. **15**(11), 0240784. <https://doi.org/10.1371/journal.pone.0240784> (2020).
53. Moreno-Pérez, O. et al. Post-acute COVID-19 syndrome incidence and risk factors: A Mediterranean cohort study. *J. Infect.* **82**(3), 378–383. <https://doi.org/10.1016/j.jinf.2021.01.004> (2021).
54. Liu, Y.-H. et al. One-year trajectory of cognitive changes in older survivors of COVID-19 in Wuhan, China: A longitudinal cohort study. *JAMA Neurol.* **79**(5), 509. <https://doi.org/10.1001/jamaneurol.2022.0461> (2022).
55. Al-Aly, Z., Xie, Y. & Bowe, B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature* **594**(7862), 259–264. <https://doi.org/10.1038/s41586-021-03553-9> (2021).
56. Al-Aly, Z., Bowe, B. & Xie, Y. Long COVID after breakthrough SARS-CoV-2 infection. *Nat. Med.* **28**(7), 1461–1467. <https://doi.org/10.1038/s41591-022-01840-0> (2022).
57. Kim, Y., Bae, S., Chang, H.-H. & Kim, S.-W. Long COVID prevalence and impact on quality of life 2 years after acute COVID-19. *Sci. Rep.* **13**(1), 11207. <https://doi.org/10.1038/s41598-023-36995-4> (2023).
58. Tsampasian, V. et al. Risk factors associated with post- COVID-19 condition: A systematic review and meta-analysis. *JAMA Intern. Med.* **183**(6), 566. <https://doi.org/10.1001/jamainternmed.2023.0750> (2023).
59. Watanabe, A., Iwagami, M., Yasuhara, J., Takagi, H. & Kuno, T. Protective effect of COVID-19 vaccination against long COVID syndrome: A systematic review and meta-analysis. *Vaccine*. **41**(11), 1783–1790. <https://doi.org/10.1016/j.vaccine.2023.02.008> (2023).
60. Ceban, F. et al. COVID-19 vaccination for the prevention and treatment of long COVID: A systematic review and meta-analysis. *Brain Behav. Immun.* **111**, 211–229. <https://doi.org/10.1016/j.bbi.2023.03.022> (2023).
61. Brannock, M. D. et al. Long COVID risk and pre-COVID vaccination in an EHR-based cohort study from the RECOVER program. *Nat. Commun.* **14**(1), 2914. <https://doi.org/10.1038/s41467-023-38388-7> (2023).
62. Trinh, N. T. et al. Effectiveness of COVID-19 vaccines to prevent long COVID: Data from Norway. *Lancet Respir. Med.* **12**(5), 33–34. [https://doi.org/10.1016/S2213-2600\(24\)00082-1](https://doi.org/10.1016/S2213-2600(24)00082-1) (2024).
63. Marra, A. R. et al. The effectiveness of COVID-19 vaccine in the prevention of post-COVID conditions: A systematic literature review and meta-analysis of the latest research. *Antimicrob. Stewardship Healthc. Epidemiol.* **3**(1), 168. <https://doi.org/10.1017/ash.2023.447> (2023).
64. Lundberg-Morris, L. et al. Covid-19 vaccine effectiveness against post-covid-19 condition among 589 722 individuals in Sweden: Population based cohort study. *BMJ*. <https://doi.org/10.1136/bmj-2023-076990> (2023).
65. Van Gils, M. J. et al. Antibody responses against SARS-CoV-2 variants induced by four different SARS-CoV-2 vaccines in health care workers in the Netherlands: A prospective cohort study. *PLoS Med.* **19**(5), 1003991. <https://doi.org/10.1371/journal.pmed.1003991> (2022).
66. Soheili, M. et al. The efficacy and effectiveness of COVID-19 vaccines around the world: A mini-review and meta-analysis. *Ann. Clin. Microbiol. Antimicrob.* **22**(1), 42. <https://doi.org/10.1186/s12941-023-00594-y> (2023).
67. Azzolini, E. et al. Association between BNT162b2 vaccination and long COVID after infections not requiring hospitalization in health care workers. *JAMA* **328**(7), 676–678. <https://doi.org/10.1001/jama.2022.11691> (2022).
68. Lundberg-Morris, L. et al. Covid-19 vaccine effectiveness against post-covid-19 condition among 589 722 individuals in Sweden: Population based cohort study. *BMJ*. <https://doi.org/10.1136/bmj-2023-076990> (2023).
69. Antony, B. et al. Predictive models of long COVID. *EBioMedicine* **96**, 104777. <https://doi.org/10.1016/j.ebiom.2023.104777> (2023).
70. Kessler, R., Philipp, J., Wilfer, J. & Kostev, K. Predictive attributes for developing long COVID—a study using machine learning and real-world data from primary care physicians in Germany. *J. Clin. Med.* **12**(10), 3511. <https://doi.org/10.3390/jcm12103511> (2023).

Author contributions

A. Guillen-Teruel contributed to the design and conceptualisation of the study, methodology, formal analysis and investigation, writing - original draft preparation, visualisation. J.L. Mellina-Andreu contributed to the web development and writing - review and editing. G. Reina and E. Gonzalez-Billalabeitia contributed to the design and conceptualization of the study, writing - review and editing. R. Rodriguez-Iborra contributed to the resources acquisition and writing - review and editing. J. Palma and J.A. Botia contributed to the design and conceptualisation of the study, writing - original draft preparation and supervision. A. Cisterna-Garcia contributed to the design and conceptualisation of the study, methodology, writing - original draft preparation and supervision.

Funding

This work was supported by the Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia (Spain) 21591/FPI/21 and 22308/FPI/23 the Spanish Council of Science and Innovation through the Grant PID2022-136306OB-I00 funded by MICIU/AEI/10.13039/501100011033 and FEDER, UE. JB was supported by the same foundation through the research project 00007/COVI/20. AC was supported by funding from European Union NextGenerationEU.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-94765-w>.

Correspondence and requests for materials should be addressed to A.C.-G.

Reprints and permissions information is available at www.nature.com/reprints.

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

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/>.

© The Author(s) 2025