Veronese et al. Top Italian Scientists Journal; 2024;1(1)
https://doi.org/10.62684/XHIY3899

Top Italian Scientists Journal

Open Access | Article | Clinical Sciences

ISSN 3033-5132

Prognostic value of clinical and microbiological parameters in COVID-19: the COMEPA study

Nicola Veronese[a], Marianna Noale[b], Anna La Carrubba[a], Luca Carruba[a], Stefano Ciriminna[a], Francesco Pollicino[a], Dario Saguto[a], Simona De Grazia[a], Federica Cacioppo[a], Giovanni M. Giammanco[a], Claudio Costantino[a], Francesco Vitale[a], Marco Affronti[c], Maria Chiara Morgante[c], Giusi Randazzo[c], Ligia J. Dominguez[a,d], Stefania Maggi[b], Mario Barbagallo[a], and the COMEPA study authors

[a] Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties "G. D'Alessandro", University of Palermo, 90127 Palermo, Italy.
[b] Consiglio Nazionale delle Ricerche, Neuroscience Institute, Padova, 35100, Italy.
[c] Geriatric Unit, Azienda Ospedaliera Universitaria Policlinico Paolo Giaccone, Palermo, Italy.
[d] School of Medicine and Surgery, University of Enna "Kore", 94100, Enna, Italy;

Correspondence to: Nicola Veronese. Geriatric Unit, Department of Internal Medicine and Geriatrics, University of Palermo, Via del Vespro, 141, 90127 Palermo, Italy. Email: nicola.veronese@unipa.it

Published: 12 Jan 2024

Abstract

Purpose

Clusters' analysis may indicate distinct phenotypes and symptom profiles potentially due to differing pathophysiology and needing different clinical approaches in COVID-19. However, the research about clusters combining clinical and microbiological information is still limited. The purpose of our study was to examine the prognostic role of clusters, including clinical and microbiological parameters in terms of severity of lung involvement, in-hospital mortality, and the occurrence of long COVID.

Methods

Information regarding COVID-19, mortality, severity of lung involvement derived from medical records; long COVID symptomatology was ascertained using phone calls. A k-means clustering method was considered to partition data into clusters considering typical symptoms of COVID-19 present at hospital admission and SarsCov2 variants.

Results

Our analysis identified among 414 patients (mean age: 65 years; males: 59.9%) four different clusters. Cluster 1: higher prevalence of respiratory COVID symptoms at hospital admission; Cluster 2: higher frequency of non-respiratory COVID symptoms and a higher prevalence of the Alpha variant; Cluster 3: older subjects and more frequently men, reporting more severe medical conditions and with a higher prevalence of Wild type variant; Cluster 4: patients that more often reported general and gastrointestinal COVID symptoms at the admission. From a prognostic point of view, patients in cluster 3 more frequently died and were admitted in a nursing home, with significantly lower presence of long COVID symptomatology.

Conclusions

Clusters combining clinical and microbiological information in individuals hospitalized with COVID-19 that had different not only different profiles, but also different prognostic values, also in terms of long COVID.

Keywords: COVID-19; clusters; long COVID; prognosis.

Introduction

Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus.[1] Early SARS-CoV-2 variants caused low respiratory tract infections, with high morbidity and mortality and symptoms evolved with the emergence of new variants.[2] Since the appearance of SARS-CoV-2 in December 2019, it has evolved more than ten variant strains.[2] Among these variants, five of them (Alpha, Beta, Gamma, Delta, and Omicron) were thought to be more transmissible and/or more lethal than the original Wuhan strain and have been designated as variants of concern by the World Health Organization (WHO).[3]

Long COVID is defined as the continuation or development of new symptoms three months after the initial SARS-CoV-2 infection, with these symptoms lasting for at least two months with no other explanation.[4] While common symptoms of long COVID can include fatigue, shortness of breath and cognitive dysfunction, over 200 different symptoms have been reported that can have an impact on everyday functioning.[5] Studies show that around 10–20% of people infected by SARS-CoV-2 may go on to develop symptoms that can be diagnosed as long COVID.[3] Although exact numbers of those living with the condition are uncertain, it is believed that more than 17 million people across the WHO European Region may have experienced it during the first two years of the pandemic (2020/21). [6]

At the same time, the interaction between vaccinations and Sars-CoV-2 variants have modified the clinical course of COVID-19 during these years of pandemic [7], indicating the possibility to have different presentations of COVID-19, i.e., clusters. Clusters' analysis may better indicate distinct phenotypes and symptom profiles potentially due to differing pathophysiology and needing different clinical approaches. Even if this kind of analysis is highly encouraged to better individualize causes and potential treatments for each patient[8], to the best of our knowledge, only a few studies have explored the importance of clusters in COVID-19 and mainly with data regarding long COVID.[9] Given this background, the purpose of our study was to examine the prognostic role of clusters including clinical and microbiological parameters in patients affected by COVID-19, in terms of severity of lung involvement, in-hospital mortality and the occurrence of long COVID.

Materials and methods

Study population

The study was conducted at the University Hospital of Palermo, which is one of the largest University Hospital in Sicily with a total of 604 hospital beds, 542 of which are ordinary beds and 62 Day Hospitals beds, 2,100 employees (including healthcare professionals and administrative staff), 1,120 medical residents and over 1,500 trainees of healthcare university courses (medicine and healthcare professions). The Palermo University Hospital is the only Hospital in Western Sicily to have an outpatient vaccination unit for the vaccination of healthcare workers, employees, hospitalized patients and the general population at high-risk (for severe comorbidities and allergic diseases).

All patients aged ≥18 years hospitalized in the Internal Medicine or Geriatrics Wards from the 1st of September 2020 to 31st May 2021 at the University Hospital 'P. Giaccone' from Palermo, Italy were enrolled.[10] No other inclusion criteria were considered to better represent a real-life scenario. The study was approved by the Local Ethical Committee during the session of the 28th of April 2021 (protocol number 04/2021). For hygienic reasons, the informed consent to participate to the study was collected orally and reported in the medical records.

Clustering assessment and phenotype profiling

A k-means clustering method, a type of unsupervised machine learning algorithm, was considered to partition data into clusters considering typical symptoms of COVID-19 present at hospital admission (fever, cough, asthenia, headache, dyspnea, anorexia, anosmia, ageusia, myalgia, arthralgia, diarrhoea, nausea, vomiting, other symptoms) and SarsCov2 variants. K-means method assigns clusters to observations in order to minimize the distance between observations and cluster centroids, allowing to obtain observations that are similar to each other in the same cluster and dissimilar to other observations in other clusters.[11] Non-overlapping clusters are defined with an iterative approach by reassigning cluster membership and cluster centroids until the solution reaches a local optimum. The optimal number of clusters was determined by user comparing pseudo F statistic and cubic clustering criterion (CCC) for models with different number of clusters (0 to 10).[12]

Outcomes

The primary outcomes were mortality during hospital stay, the presence of long COVID during the follow-up, and severity of lung involvement. Information regarding mortality was collected using clinical records and death certificates. Long COVID-19 symptomatology was defined using the World Health Organization (WHO) indications, i.e., "*a condition that occurs in individuals with a history of probable or confirmed SARS-CoV-2 infection, usually 3 months from the onset of COVID-19 with symptoms that last for at least 2 months and cannot be explained by an alternative diagnosis*". [4] Accordingly, the presence of long COVID was assessed after a median of 17 months (range: 13-22) from hospital discharge through phone calls, using a method largely made in other works.[13-16] We considered as signs or symptoms of long COVID those indicated in several systematic reviews [13-16], i.e., neurological, respiratory, mobility impairment, heart, digestive, skin, or general signs and symptoms that can be attributable to COVID-19 infection. All the questions were posed as yes/no questions by phone. Psychiatric conditions, considered as secondary outcomes of the work, were assessed using the Post-traumatic Stress Disorder (PTSD) Checklist (PCL)-5 [17] and the Hospital Anxiety and Depression Scale (HADS), a validated tool for evaluating anxiety and depression among adults.[18] Respiratory failure was defined as a partial pressure of oxygen (PaO2) < 60 mmHg with a normal or decreased partial pressure of carbon dioxide (PaCO2).[19] The severity of lung involvement was made according to the Berlin Criteria[20] in mild ($200 > PaO2/FiO2 \leq 300$), moderate ($100 > PaO2/FiO2 \leq 200$) or severe ($PaO2/FiO2 \leq 100$).

Covariates

Among the parameters collected in the COMEPA study [10], for the aim of the present study we used the information potentially affecting the association between clusters and the outcomes of interest mentioned before. Therefore, we included as factors: age; gender; comorbidities evaluated in terms of presence and the severity using the Cumulative Illness Rating Scale (CIRS) [21] that estimates the severity of pathology in each of 13 systems, with a grade from 0 to 4, with a value $\geq$ 2 indicating the presence of, at least, a moderate disease.[21] Moreover, the presence of signs or symptoms typical of COVID-19 and present at hospital admission was recorded using medical records, physical examination, and recent medical history.

COVID-19 Vaccination history was recorded using the National Vaccination Registry (NVR). The COVID-19 NVR started in February 2021 and allowed to all Health care professionals that administered COVID-19 vaccines to general population to charge data on a web portal through personal Username and password (https://www.governo.it/it/cscovid19/report-vaccini/; last access 10th of March 2023).

COVID-19 variants were ascertained using the SARS-CoV-2 Extended ELITe MGB® Kit able to detect and discriminate the mutations L452R, E484K, E484Q and N501Y of the S gene of SARS-CoV-2 through Reverse Transcription and Real-Time Polymerase Chain Reaction and melting curve analysis. Finally, the length of stay in hospital was also included as covariate in the analysis.

Statistical analysis

Participants' characteristics are presented as counts and percentages for categorical variables, and as means ± standard deviation (SD) for quantitative measures. No imputation of missing values was made. Comparisons were performed considering the Chi-squared or the Fisher exact tests for categorical variables or generalized linear model testing for homoschedasticity (Levene's test) for quantitative variables.

A cluster analysis based on k-means method was applied to classify participants into clusters considering their symptoms. Characteristics of participants in different clusters were compared considering Chi-squared, Fisher exact tests or generalized linear model, as appropriate. Associations between clusters and different outcomes (severity of lung involvement, dichotomized into no vs mild, moderate or severe; long COVID symptoms; death during follow-up) were evaluated considering multivariable logistic regression models adjusted for age, sex, comorbidity and duration of the hospitalization, and results were presented as odds ratio (OR) and 95% confidence interval (CI). In relation to post-traumatic stress disorder score according to PTSD and to anxiety/depression score according to HADS, generalized linear models with a Poisson distribution family and a log link function were considered.

All statistical tests were two-tailed and statistical significance was assumed for p-value <0.05. The analyses were performed using SAS, V.9.4 (SAS Institute, Cary, NC).

Results

Among 530 patients initially included in the COMEPA study, 414 were analyzed since, for 116 individuals, no sufficient information was available for the aims of this study. Overall, as shown in **Supplementary Table 1**, the patients aged a mean of 65.0±15.2 years and they were prevalently males (59.9%). Over a median of 17 months (range: 13–22) from hospital discharge, 115 (=27.8%) reported a long COVID symptomatology. **Supplementary Figure 1** reports the prevalence of the single long COVID signs and symptoms, overall showing that the most frequent was weakness.
Overall, 6 out 414 patients (coverage rate: 1,4%) were vaccinated against COVID-19 before hospitalization, and the impact of vaccination was not significant on our clusters (data not shown in table).

Patients having long COVID were significantly younger ($p<0.001$) than their counterparts. Regarding medical conditions, people with long COVID had, during hospitalization, a lower prevalence of cardiac ($p=0.003$) and renal ($p=0.019$) conditions. No significant differences emerged in terms of COVID-19 variants (**Supplementary Table 1)**. Patients reporting long COVID during follow-up had a significantly higher prevalence of COVID symptoms when admitted ($p=0.030$), in particular fever, asthenia, arthralgia ($p<0.05$ for all these comparisons). No different regarding the severity of lung involvement was observed, whilst patients reporting long COVID were discharged more frequently at home **(Supplementary Material 1**).

Table 1 shows the baseline characteristics by clusters. Briefly, Cluster 1 was composed by people who tend to be older, who have had cough or dyspnea as COVID symptoms at hospital admission, with frequent moderate or severe respiratory failure, discharged especially at home or in nursing home. Cluster 2 included subjects that tend to be younger with a higher prevalence of the Alpha variant. In cluster 2, the diseases at the admission were less frequent and less serious according to the CIRS. In cluster 2, COVID symptoms at the hospital admission included fever, cough, asthenia, anosmia, ageusia, myalgia and arthralgia, with frequent moderate or severe respiratory failure during the hospitalization. Patients included in cluster 3 tend to be older and more frequently men, reporting more severe medical conditions according to the CIRS. The wild variant was more frequent in this cluster. In cluster 3, patients had fewer COVID symptoms after hospital admission, but reported more often altered mental status. In cluster 3, the discharge to nursing

home or death following hospitalization were more frequent. Finally, cluster 4 included patients that more often reported other COVID symptoms at the admission, such as abdominal or chest pain, dizziness, or syncope, even if they reported less frequently respiratory failure. Overall, regarding long COVID, cluster 1 reported higher values of HADS; cluster 2 had the highest presence of long COVID, such as weakness and hair loss, and reaching high HADS scores.

Table 2 shows the association between the four clusters and the incidence of long COVID, death and severity of lung involvement. After adjusting the analyses for age, sex, severity of medical conditions, according to the CIRS, and duration of the hospitalization, patients in cluster 3 reported a significantly lower presence of long COVID symptomatology (OR=0.40; 95%CI: 0.17-0.96; p=0.041) or severity of lung involvement (OR=0.44; 95%CI:0.21-0.93; p=0.031) than patients in cluster 2. On the contrary no significant differences emerged for mortality across the clusters (**Table 2**). Similarly, when considering psychiatric conditions during follow-up as outcome, patients in cluster 4 reported a significantly lower presence of PTSD than in cluster 2 (RR=0.41: 95%CI: 0.19-0.92; p=0.030), whilst subjects in cluster 1 reported more frequently than cluster 2 evidence of depression/anxiety (RR=1.57; 95%CI: 1.29-1.92; p<0.001) (**Table 3**).

Discussion

In our study, we profiled phenotypes of hospitalized patients affected by COVID-19, overall identifying four different symptom profiles (or clusters) which varied in number and symptom combinations. Altogether these findings may indicate the existence of heterogeneous profiles that finally lead to a different prognosis. The main clusters of our analysis were composed by one group of people with a higher prevalence of respiratory COVID symptoms at hospital admission; another group with a higher frequency of non-respiratory COVID symptoms and a higher prevalence of the Alpha variant; a third group with subjects older and more frequently men, reporting more severe medical conditions and with a higher prevalence of Wild type variant. A last cluster included patients that more often reported general and gastrointestinal COVID symptoms at the admission. In our work, we reported that these clusters have a different impact on the outcomes of our research, such as mortality, respiratory failure and long COVID symptomatology. The cluster are graphically reported in **Figure 1**.

Regarding the importance of identifying clusters having different prognostic role, we can mention other important studies. First, the Post-hospitalization COVID-19 (PHOSP) study that evaluated long-term symptoms in hospitalized patients, that identified four clusters with different physical and health impairment profiles.[22] However, this study did not incorporate the role of SarsCov2 variants and COVID symptoms. At the same, a study from the National Core Study for Health and Wellbeing incorporated individuals from the general population identifying two clusters in individuals more than 12 weeks after SarSCov2 infection, i.e., high and low symptom burden clusters.[23] Finally, another most recent work, made in the general population using a phone app, reported the presence of three main clusters, i.e., one cluster dominated by central neurological symptoms, a second cluster dominated by cardiorespiratory symptoms, and a third more heterogeneous cluster showing systemic and inflammatory symptoms. [7]

We believe that our study completes these important findings since other factors were explored in the clusters and we added some outcomes of clinical importance such as in-hospital mortality, respiratory failure during hospitalization, and long COVID symptomatology. In detail, patients included in cluster 3 (older, more frequently men, with more severe chronic medical conditions, and with a higher prevalence of Wild variant) died more frequently and were admitted more

frequently in a nursing home. At the same time, the patients included in the cluster 3 reported a significantly lower presence of long COVID symptomatology than other patients and respiratory failure during hospitalization. Overall, these findings that can be ambiguous may suggest that older people died more frequently for acute complications of medical chronic conditions that they had before hospitalization that for COVID-19 *per se*.[24] At the same time, patients in cluster 2, i.e., younger patients with a higher frequency of several COVID symptoms at hospital admission, had a significantly higher prevalence of English type variant, confirming the findings of the previous study reporting this kind of association.[25] Of importance, this cluster was associated with a significantly higher presence of long COVID symptomatology that is, in our opinion, of public health importance since these patients were the youngest in age among those considered in our analyses and overall confirming that the presence of long COVID is of particular importance in adults more than in older people.[5] When considering psychiatric conditions during follow-up, patients in cluster 1 (i.e., old subjects with a higher prevalence of cough and dyspnea at hospital admission) had, more frequently, respiratory failure during hospitalization. Of importance, this cluster reported a higher prevalence of depression/anxiety during follow-up. Maybe, we can justify this finding due to the use of invasive and non-invasive respiratory supports that seem to be associated with unfavorable mental health outcomes. [26,27]

We also analyze vaccination status against COVID-19 in these clusters and the possible impact on clinical outcomes. Unfortunately, only six patients over 414 were vaccinated before the hospitalization, probably because during the enrollment period (from 1st of September 2020 to 31st May 2021) COVID-19 vaccination rate among general population were still low among general population under 80 years old. Vaccination of HCPs started in January 2021, vaccination of the elderly and of extremely vulnerable patients started at the end of February 2021 and, due to limited doses of mRNA vaccines, at least 50% of subjects older than 60 y.o. received in Italy only the first dose of adenovirus vector vaccines against COVID since the end of March 2021.
(Ref: Istituto Superiore di Sanità - EpiCentro - Epidemiology for public health. National COVID-19 vaccination plan. Available online: https://www.epicentro.iss.it/en/vaccines/covid-19-vaccination-plan). At the same time, vaccination is important not only because lowers mortality [28], but also because it seems to be associated with a lower risk of long COVID [29] and therefore could significantly modify the associations that we found.

The findings of our study must be interpreted within its limitations. First, only hospitalized people were included probably introducing a selection bias in our findings. Second, as mentioned before, the important role of vaccinations was not explored due to a limited number of patients vaccinated before the hospitalization. Lastly, our model only used the symptoms assessed by our daily clinical practice and we did not consider symptoms reported as free text; thus, we might ignore other symptoms that could affect the clusters.

In conclusion, we identified clusters combining clinical and microbiological information in individuals hospitalized with COVID-19 that had not only different profiles, but also different prognostic values. We believe that our findings may have relevance to individuals previously affected by COVID-19, their general practitioners, and in a public health perspective, providing other information to validate their illness and manage their expectations. Our findings may add to the emerging evidence that long COVID may have sub-types, possibly with differing pathophysiology.[7] Further investigation also using artificial intelligence models can better integrate not only clinical information, but also others (such as laboratory, radiological and genetic) that can further explain the reasons of different prognostic profiles, also in terms of long COVID symptomatology.

Declarations

Acknowledgements

The COMEPA group includes (alphabetical order): Affronti Marco, Amodeo Simona, Barbagallo Mario, Briganò Vincenza Maria, Cacioppo Federica, Capitano Walter M., Carruba Luca, Cavaleri Francesco, Catanese Giuseppina, Citarrella Roberto, Di Bella Giovanna, Di Franco Giuseppina, Di Prazza Agnese, Dominguez Ligia Juliana, Giannitrapani Lydia, Grasso Giulia, Immordino Federico, Licata Anna, La Carruba Anna, Mansueto Pasquale, Mirarchi Luigi, Morgante Maria Chiara, Parinello Alessandra, Pecoraro Emanuela, Peralta Marco, Polizzotto Carla, Pollicino Francesco, Quartetti Federico, Randazzo Giusi, Rizzo Angelo, Rizzo Giuseppina, Sanfilippo Valeria, Soresi Maurizio, Malerba Valentina, Vernuccio Laura, Veronese Nicola, Zerbo Maddalena.

Conflict of interest

The Authors declare that there is no conflict of interest.

Funding

None.

Data Availability

The datasets analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval

This study was performed in line with the principles of the Declaration of Helsinki. The study was approved by the Local Ethical Committee during the session of the 28th of April 2021 (protocol number 04/2021).

Consent to participate

Informed consent was obtained from all individual participants included in the study.

FIGURE LEGEND

Figure 1. Clusters of clinical and microbiological parameters in the COMEPA study.

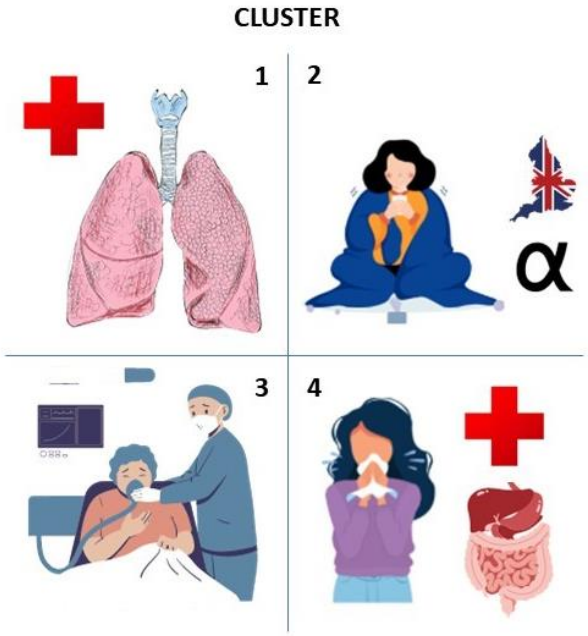


Table 1. Baseline participants' characteristics by clusters

	Cluster 1 (n=208)	Cluster 2 (n=92)	Cluster 3 (n=102)	Cluster 4 (n=93)	p-value
Sex, female, n (%)	87 (41.8)	44 (47.8)	40 (39.2)	45 (48.4)	0.458
Age, years, mean±SD	65.6±13.9	62.5±14.9	66.6±16.7	64.1±16.3	0.109
Diseases with moderate, severe or extremely severe problems, n (%)					
Cardiac (heart only)	30 (19.6)	11 (15.1)	23 (31.9)	11 (19.3)	0.072
Vascular	78 (51.0)	36 (48.0)	35 (48.6)	23 (40.4)	0.596
Hematological	6 (4.0)	9 (12.0)	16 (22.2)	6 (10.5)	0.001
Respiratory	18 (11.8)	2 (2.7)	11 (15.5)	8 (14.0)	0.063
Ophthalmological, otorhinolaryngology	4 (2.6)	3 (4.1)	2 (2.8)	3 (5.3)	0.762

	Cluster 1 (n=208)	Cluster 2 (n=92)	Cluster 3 (n=102)	Cluster 4 (n=93)	p-value
Upper gastrointestinal	3 (2.0)	1 (1.4)	7 (9.7)	4 (7.0)	0.016
Lower gastrointestinal	5 (3.3)	6 (8.1)	5 (6.9)	8 (14.0)	0.042
Hepatic and pancreatic	6 (3.9)	3 (4.1)	9 (12.5)	6 (10.5)	0.047
Renal	12 (7.8)	5 (6.8)	7 (9.7)	4 (7.0)	0.913
Genitourinary	8 (5.2)	1 (1.4)	6 (8.3)	3 (5.3)	0.267
Musculoskeletal and tegumental	12 (7.8)	4 (5.4)	5 (6.9)	2 (3.5)	0.727
Neurological	14 (9.2)	4 (5.4)	12 (16.7)	6 (10.5)	0.146
Endocrine, metabolic	37 (24.2)	23 (31.1)	23 (31.9)	14 (24.6)	0.516
Psychiatric	12 (7.8)	1 (1.4)	5 (6.9)	1 (1.8)	0.107
CIRS Comorbidity Index, mean±SD	0.64±1.42	0.51±0.91	1.19±1.27	0.69±1.64	<0.001
CIRS Severity Index, mean±SD	0.32±0.28	0.28±0.29	0.52±0.39	0.36±0.30	<0.001
COVID-19 variant, n (%)					0.029
Wild	119 (72.6)	48 (60.0)	68 (81.0)	56 (70.0)	
English	45 (27.4)	32 (40.0)	16 (19.0)	24 (30.0)	
Alterated mental state when admitted, n (%)	18 (8.7)	1 (1.1)	15 (14.9)	10 (10.8)	0.008
COVID symptoms when admitted, n (%)	208 (100.0)	92 (100.0)	66 (64.7)	93 (100.0)	<0.001
Fever, n (%)	148 (71.2)	84 (91.3)	3 (2.9)	20 (21.5)	<0.001
Cough, n (%)	68 (32.7)	49 (53.3)	7 (6.9)	9 (9.7)	<0.001
Asthenia, n (%)	23 (11.1)	73 (79.4)	17 (16.7)	15 (16.1)	<0.001
Headache, n (%)	2 (1.0)	15 (16.3)	11 (10.8)	5 (5.4)	<0.001
Dyspnea, n (%)	159 (76.4)	30 (32.6)	0 (0.0)	14 (15.1)	<0.001
Anorexia, n (%)	6 (2.9)	3 (3.3)	8 (7.8)	2 (2.2)	0.171
Anosmia, n (%)	8 (3.9)	13 (14.1)	4 (3.9)	3 (3.2)	0.002
Ageusia, n (%)	5 (2.4)	11 (12.0)	2 (2.0)	4 (4.3)	0.004
Myalgia, n (%)	4 (1.9)	26 (28.3)	8 (7.8)	5 (5.4)	<0.001
Arthralgia, n (%)	7 (3.4)	35 (38.0)	6 (5.9)	3 (3.2)	<0.001
Diarrhea, n (%)	15 (7.2)	15 (16.3)	8 (7.8)	6 (6.5)	0.051
Nausea, n (%)	3 (1.4)	3 (3.3)	4 (3.9)	6 (6.5)	0.126

	Cluster 1 (n=208)	Cluster 2 (n=92)	Cluster 3 (n=102)	Cluster 4 (n=93)	p-value
Vomiting, n (%)	5 (2.4)	6 (6.5)	7 (6.9)	6 (6.5)	0.137
Other symptoms, n (%)	7 (3.4)	7 (7.6)	0 (0.0)	93 (100.0)	<0.001
Abdominal pain	0 (0.0)	4 (4.4)	0 (0.0)	25 (26.9)	<0.001
Chest pain	4 (1.9)	1 (1.1)	0 (0.0)	25 (26.9)	<0.001
Vertigo/lipothymia	2 (1.0)	2 (2.2)	0 (0.0)	24 (25.8)	<0.001
Syncope	0 (0.0)	0 (0.0)	0 (0.0)	7 (7.5)	<0.001
Gastrointestinal bleeding	0 (0.0)	0 (0.0)	0 (0.0)	3 (3.2)	0.005
Faringodinia	1 (4.8)	0 (0.0)	0 (0.0)	5 (5.4)	0.001
Hematuria	0 (0.0)	0 (0.0)	0 (0.0)	4 (4.3)	0.001
Respiratory failure, n (%)	14 (6.7)	9 (9.8)	7 (6.9)	3 (3.2)	0.360
Severity of lung involvement, n (%)					0.004
None	117 (56.5)	56 (60.9)	80 (78.4)	67 (72.0)	
Mild	60 (29.0)	19 (20.7)	13 (12.8)	16 (17.2)	
Moderate	21 (10.1)	14 (15.2)	9 (8.8)	9 (9.7)	
Severe	9 (4.4)	3 (3.3)	0 (0.0)	1 (1.1)	
N. of symptoms when admitted, mean±SD	2.3±1.0	4.1±1.4	1.2±1.2	2.1±1.0	<0.001
Duration of the hospitalization, days, mean±SD	13.6±9.6	15.5±26.7	14.8±13.7	12.5±10.2	0.719
Type of hospital discharge, n (%)					0.031
Home	127 (68.8)	66 (71.7)	53 (52.0)	63 (67.7)	
Voluntary	5 (2.4)	0 (0.0)	3 (2.9)	3 (3.2)	
Institute	35 (16.8)	12 (13.0)	18 (17.7)	18 (19.4)	
Other Hospital ward	8 (3.9)	7 (7.6)	6 (5.9)	2 (2.2)	
Intensive care	6 (2.9)	1 (1.1)	4 (3.9)	0 (0.0)	
Death	11 (5.3)	6 (6.5)	17 (16.7)	7 (7.5)	
Hospice	0 (0.0)	0 (0.0)	1 (1.0)	0 (0.0)	

Abbreviations: CIRS (Cumulative Illness Rating Scale; SD (Standard Deviation)

Table 2. Association between clusters and different outcomes (logistic regression models)

	Long COVID			Death			Severity of lung involvement (mild, moderate or severe)		
	OR	95% CI	p-value	OR	95% CI	p-value	OR	95% CI	p-value
Cluster 1 vs 2	0.68	0.34-1.35	0.272	1.73	0.52-5.77	0.375	1.38	0.77-2.47	0.282
Cluster 3 vs 2	0.40	0.17-0.96	0.041	3.43	0.89-13.2	0.073	0.44	0.21-0.93	0.031
Cluster 4 vs 2	0.66	0.28-1.58	0.349	1.13	0.24-5.26	0.881	0.63	0.30-1.36	0.240
Sex, females vs males	1.02	0.59-1.78	0.940	2.56	1.03-6.39	0.044	0.75	0.47-1.20	0.234
Age, years	0.97	0.95-0.99	0.004	1.12	1.07-1.17	<0.001	1.02	1.01-1.04	0.013
CIRS Comorbidity Index	0.68	0.49-0.95	0.024	1.67	1.19-2.34	0.003	1.01	0.85-1.19	0.972
Duration of the hospitalization	1.02	0.99-1.04	0.121	1.03	0.99-1.06	0.106	1.02	0.99-1.04	0.110

Abbreviations: CI (Confidence Interval); OR (Odds Ratio)

Table 3. Clusters and post-traumatic stress disorder, and anxiety/depression (generalized linear models with a Poisson distribution family and log link function)

	PTSD			HADS		
	RR	95% CI	p-value	RR	95% CI	p-value
Cluster 1 vs 2	0.90	0.16-1.75	0.753	1.57	1.29-1.92	<0.001
Cluster 3 vs 2	0.60	0.27-1.32	0.205	1.11	0.86-1.42	0.416
Cluster 4 vs 2	0.41	0.19-0.92	0.030	1.16	0.91-1.46	0.230
Sex, females vs males	1.38	0.89-2.15	0.152	1.28	1.12-1.47	<0.001
Age, 5 years	1.00	0.92-1.09	0.929	0.96	0.94-0.98	0.001
CIRS Comorbidity Index	1.02	0.97-1.08	0.350	0.90	0.84-0.96	0.001
Duration of the hospitalization	1.08	0.95-1.24	0.228	1.11	1.08-1.14	<0.001

Supplementary Table 1. Baseline characteristics of the cohort by long COVID symptoms reported.

		Long COVID		
	Overall (n=414)	No (n=299)	Yes (n=115)	p-value
Sex, female, n (%)	170 (41.1)	121 (40.5)	49 (42.6)	0.692
Age, years, mean±SD	65.0±15.2	67.1±15.5	59.8±13.0	<0.001
Diseases with moderate, severe or extremely severe problems, n (%)				
Cardiac (heart only)	64 (21.4)	56 (25.8)	8 (9.8)	0.003
Vascular	144 (47.8)	102 (46.6)	42 (51.2)	0.473
Hematological	35 (11.7)	29 (13.3)	6 (7.3)	0.150
Respiratory	34 (11.4)	28 (12.9)	6 (7.3)	0.175
Ophthalmological and otorhinolaryngology	9 (3.0)	7 (3.2)	2 (2.4)	1.000
Upper gastrointestinal	10 (3.3)	8 (3.7)	2 (2.4)	0.597
Lower gastrointestinal	21 (7.0)	13 (6.0)	8 (9.8)	0.251
Hepatic and pancreatic	20 (6.7)	17 (7.8)	3 (3.7)	0.200
Renal	26 (8.7)	24 (11.0)	2 (2.4)	0.019
Genitourinary	16 (5.3)	13 (6.0)	3 (3.7)	0.570
Musculoskeletal and tegumental	22 (7.3)	17 (7.8)	5 (6.1)	0.615
Neurological	31 (10.3)	27 (12.4)	4 (4.9)	0.057
Endocrine, metabolic	81 (27.0)	61 (28.0)	20 (24.4)	0.532
Psychiatric	15 (5.0)	14 (6.4)	1 (1.2)	0.077
CIRS Comorbidity Index, mean±SD	0.76±1.31	0.91±1.41	0.37±0.88	<0.001
CIRS Severity Index, mean±SD	0.36±0.33	0.40±0.35	0.26±0.24	0.003
COVID-19 variant, n (%)				0.940
Wild	240 (71.4)	169 (71.3)	71 (71.7)	
English	96 (28.6)	68 (28.7)	28 (28.3)	

	Overall (n=414)	Long COVID		p-value
		No (n=299)	Yes (n=115)	
Alterated mental state when admitted, n (%)	40 (9.7)	36 (12.1)	4 (3.5)	0.008
COVID symptoms when admitted, n (%)	380 (91.8)	269 (90.0)	111 (96.5)	0.030
Fever, n (%)	210 (50.7)	139 (46.5)	71 (61.7)	0.005
Cough, n (%)	109 (26.3)	71 (23.8)	38 (33.0)	0.054
Asthenia, n (%)	105 (25.4)	65 (21.7)	40 (34.8)	0.006
Headache, n (%)	22 (5.3)	12 (4.0)	10 (8.7)	0.057
Dyspnea, n (%)	169 (40.8)	121 (40.5)	48 (41.7)	0.814
Anorexia, n (%)	17 (4.1)	15 (5.0)	2 (1.7)	0.132
Anosmia, n (%)	21 (5.1)	15 (5.0)	6 (5.2)	0.934
Ageusia, n (%)	14 (3.4)	8 (2.7)	6 (5.2)	0.227
Myalgia, n (%)	39 (9.4)	26 (8.7)	13 (11.3)	0.416
Arthralgia, n (%)	45 (10.9)	26 (8.7)	19 (16.5)	0.022
Diarrhea, n (%)	38 (9.2)	25 (8.4)	13 (11.3)	0.353
Nausea, n (%)	11 (2.7)	9 (3.0)	2 (1.7)	0.735
Vomiting, n (%)	17 (4.1)	12 (4.0)	5 (4.4)	1.000
Abdominal pain	25 (6.0)	17 (5.7)	8 (7.0)	0.627
Chest pain	24 (5.8)	16 (5.4)	8 (7.0)	0.531
Vertigo/lipothymia	18 (4.4)	12 (4.0)	6 (5.2)	0.591
Syncope	7 (1.7)	6 (2.0)	1 (0.9)	0.679
Gastrointestinal bleeding	2 (0.5)	2 (0.7)	0 (0.0)	1.000
Faringodinia	4 (1.0)	1 (0.3)	3 (2.6)	0.067
Hematuria	4 (1.0)	4 (1.3)	0 (0.0)	0.213
Others	47 (11.4)	35 (11.7)	12 (10.4)	0.715
Respiratory failure, n (%)	31 (7.5)	19 (6.4)	12 (10.4)	0.161
Severity of lung involvement, n (%)				0.679
None	259 (62.7)	192 (64.4)	67 (58.3)	
Mild	93 (22.5)	65 (21.8)	28 (24.3)	
Moderate	48 (11.6)	32 (10.7)	16 (13.9)	

	Overall (n=414)	Long COVID		p-value
		No (n=299)	Yes (n=115)	
Severe	13 (3.2)	9 (3.0)	4 (3.5)	
Duration of the hospitalization, days, mean±SD	14.1±16.1	14.4±17.8	13.6±10.8	0.467
Type of hospital discharge, n (%)				<0.001
Home	264 (63.8)	163 (54.5)	101 (87.8)	
Voluntary	11 (2.7)	10 (3.3)	1 (0.9)	
Institute	65 (15.7)	54 (18.1)	11 (9.6)	
Other Hospital ward	21 (5.1)	19 (6.4)	2 (1.7)	
Intensive care	11 (2.7)	11 (3.7)	0 (0.0)	
Death	41 (9.9)	41 (13.7)	0 (0.0)	
Hospice	1 (0.2)	1 (0.3)	0 (0.0)	

Abbreviations: CIRS (Cumulative Illness Rating Scale); SD (Standard Deviation)

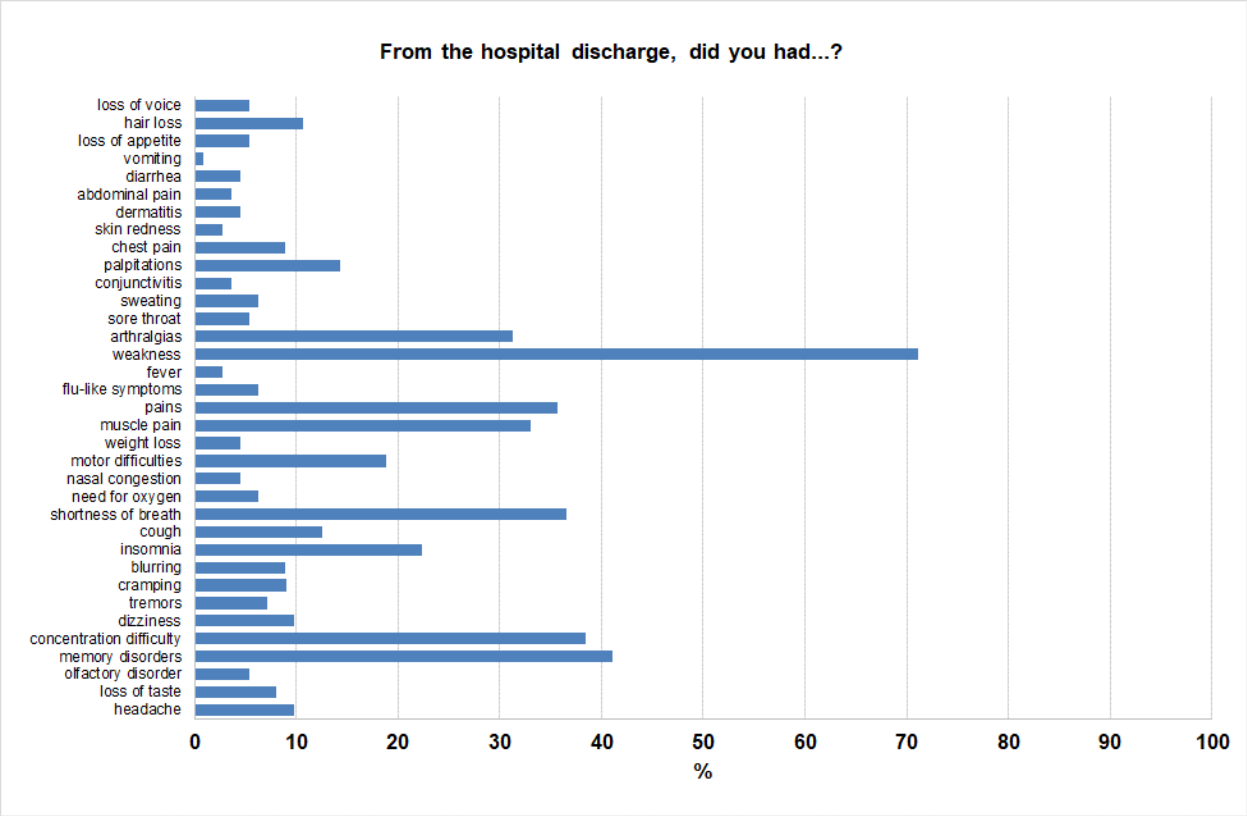
Supplementary Table 2. Long COVID characteristics by clusters

	Cluster 1 (n=208)	Cluster 2 (n=92)	Cluster 3 (n=102)	Cluster 4 (n=93)	p-value
From the hospital discharge, did you have…, n (%)					
headache	3 (3.8)	4 (8.2)	2 (8.3)	2 (4.9)	0.645
loss of taste	2 (2.5)	1 (2.0)	2 (8.3)	4 (9.8)	0.160
olfactory disorder	3 (3.8)	1 (2.0)	0 (0.0)	2 (4.9)	0.831
memory disorders	17 (21.5)	15 (30.6)	5 (20.8)	9 (22.0)	0.671
concentration difficulty	16 (20.3)	12 (24.5)	7 (29.2)	8 (19.5)	0.740
dizziness	5 (6.3)	1 (2.0)	3 (12.5)	2 (4.9)	0.338
tremors	4 (5.1)	2 (4.1)	1 (4.2)	1 (2.4)	0.956
cramping	3 (3.9)	5 (10.2)	0 (0.0)	2 (4.9)	0.328
blurring	5 (6.3)	2 (4.1)	2 (8.3)	1 (2.4)	0.702
insomnia	10 (12.7)	7 (14.3)	3 (12.5)	5 (12.2)	0.991
cough	2 (2.5)	3 (6.1)	5 (20.8)	4 (9.8)	0.023
shortness of breath	18 (22.8)	12 (24.5)	4 (16.7)	7 (17.1)	0.766
need for oxygen	3 (3.9)	2 (4.1)	1 (4.2)	1 (2.4)	1.000
nasal congestion	3 (3.8)	2 (4.1)	0 (0.0)	0 (0.0)	0.726
motor difficulties	9 (11.4)	6 (12.2)	3 (12.5)	3 (7.3)	0.873
weight loss	3 (3.8)	2 (4.1)	0 (0.0)	0 (0.0)	0.726
muscle pain	13 (16.5)	12 (24.5)	7 (29.2)	5 (12.2)	0.250
pains	14 (18.0)	12 (24.5)	6 (25.0)	8 (19.5)	0.780
flu-like symptoms	2 (2.5)	3 (6.3)	1 (4.2)	1 (2.4)	0.672
fever	0 (0.0)	1 (2.1)	1 (4.0)	1 (2.4)	0.224
weakness	35 (43.2)	25 (51.0)	10 (41.7)	11 (26.8)	0.135
arthralgia	10 (12.7)	11 (22.5)	6 (25.0)	8 (19.5)	0.388
sore throat	4 (5.1)	1 (2.0)	0 (0.0)	1 (2.4)	0.784
sweating	2 (2.5)	4 (8.2)	0 (0.0)	1 (2.4)	0.363
conjunctivitis	1 (1.3)	1 (2.0)	1 (4.2)	1 (2.4)	0.820
palpitations	9 (11.4)	4 (8.2)	2 (8.3)	1 (2.4)	0.423

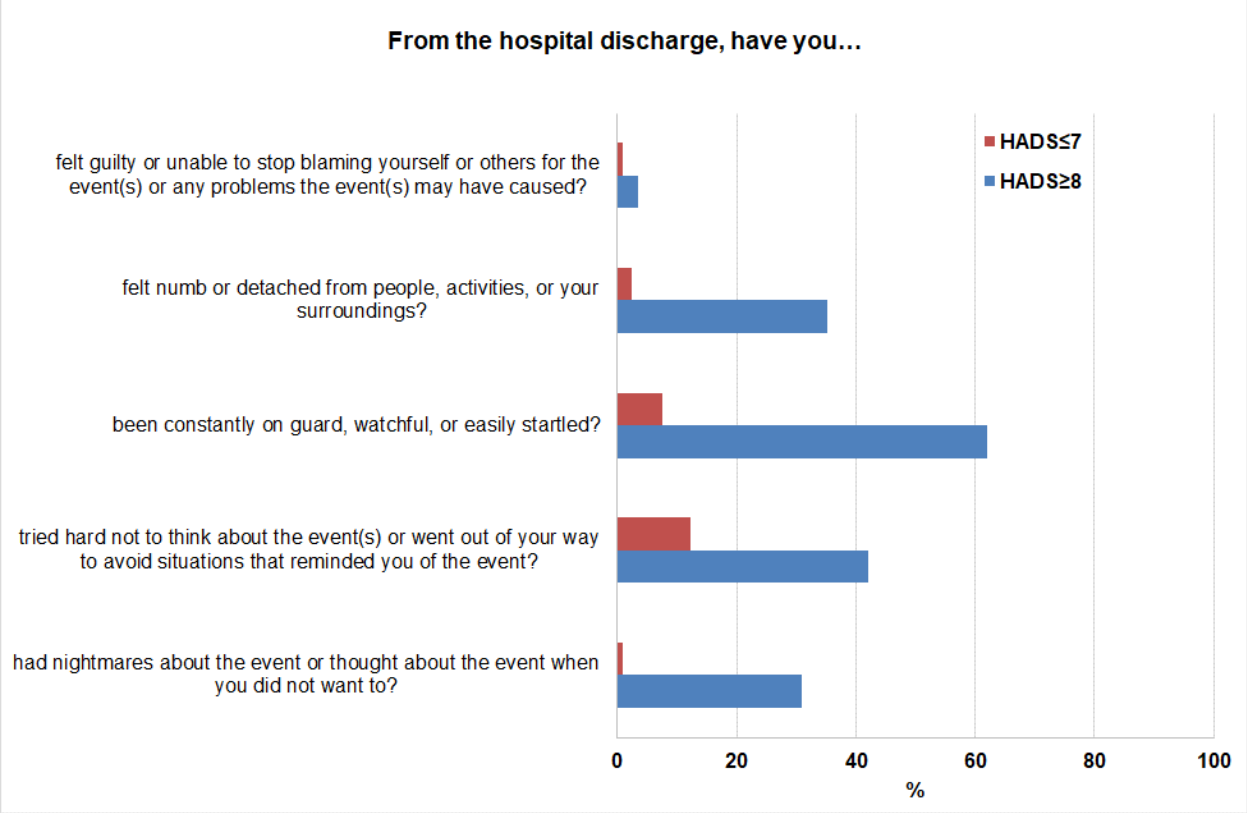
	Cluster 1 (n=208)	Cluster 2 (n=92)	Cluster 3 (n=102)	Cluster 4 (n=93)	p-value
chest pain	3 (3.8)	2 (4.1)	2 (8.3)	3 (7.3(	0.628
skin redness	2 (2.5)	1 (2.0)	0 (0.0)	0 (0.0)	0.866
dermatitis	3 (3.8)	1 (2.0)	0 (0.0)	1 (2.4)	1.000
abdominal pain	1 (1.3)	3 (6.1)	0 (0.0)	0 (0.0)	0.213
diarrhea	1 (1.3)	2 (4.1)	1 (4.2)	1 (2.4)	0.638
vomiting	1 (1.3)	0 (0.0)	0 (0.0)	0 (0.0)	1.000
loss of appetite	2 (2.5)	1 (2.0)	1 (4.2)	2 (4.9)	0.738
hair loss	5 (6.3)	7 (14.3)	0 (0.0)	0 (0.0)	0.023
loss of voice	4 (5.1)	1 (2.0)	1 (4.2)	0 (0.0)	0.552
At least one long COVID symptoms, n (%)	48 (27.4)	32 (42.7)	16 (17.2)	19 (26.8)	0.004
N. of long COVID symptoms, mean±SD	2.7±3.7	3.4±4.0	3.1±3.2	2.3±3.8	0.294
From the hospital discharge, did you, n (%)					
had nightmares about the event or thought about the event when you did not want to?	7 (9.0)	5 (10.2)	5 (21.7)	2 (5.0)	0.218
tried hard not to think about the event(s) or went out of your way to avoid situations that reminded you of the event?	17 (22.1)	9 (18.4)	7 (31.8)	7 (17.5)	0.557
been constantly on guard, watchful, or easily startled?	23 (29.1)	7 (14.3)	8 (34.8)	8 (19.5)	0.133
felt numb or detached from people, activities, or your surroundings?	11 (14.5)	5 (10.2)	3 (13.0)	4 (10.3)	0.893
felt guilty or unable to stop blaming yourself or others for the event(s) or any problems the event(s) may have caused?	2 (2.6)	1 (2.0)	0 (0.0)	0 (0.0)	0.865
PTSD, mean±SD	0.8±1.2	0.6±1.1	1.0±1.4	0.5±1.0	0.470
PTSD≥3, n (%)	10 (13.2)	4 (8.2)	5 (22.7)	3 (7.7)	0.306
HADS, mean±SD	2.8±5.9	2.9±4.5	1.8±5.1	2.0±4.3	<0.001
HADS≥8, n (%)	27 (13.0)	12 (13.0)	10 (9.8)	10 (10.8)	0.828

Abbreviations: CIRS (Cumulative Illness Rating Scale); HADS (Hospital Anxiety and Depression Scale); PTSD (Post Traumatic Stress Disorder); SD (Standard Deviation)

Supplementary Figure 1. Prevalence of long COVID characteristics.



Supplementary Figure 2. Prevalence of post-traumatic stress disorder and anxiety/depression



21/107 patients (19.6%) with long-COVID had post-traumatic stress disorder (PTSD≥3)

46/96 patients (47.9%) had a score≥8 at the HADS.

References

1. Park, M.; Cook, A.R.; Lim, J.T.; Sun, Y.; Dickens, B.L. A systematic review of COVID-19 epidemiology based on current evidence. *J. Clin. Med.* **2020**, *9*, 967.
2. Vasireddy, D.; Vanaparthy, R.; Mohan, G.; Malayala, S.V.; Atluri, P. Review of COVID-19 variants and COVID-19 vaccine efficacy: what the clinician should know? *Journal of Clinical Medicine Research* **2021**, *13*, 317.
3. Parums, D.V. Revised World Health Organization (WHO) terminology for variants of concern and variants of interest of SARS-CoV-2. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research* **2021**, *27*, e933622-933621.
4. Organization, W.H. *A clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021*; World Health Organization: 2021.
5. Di Gennaro, F.; Belati, A.; Tulone, O.; Diella, L.; Fiore Bavaro, D.; Bonica, R.; Genna, V.; Smith, L.; Trott, M.; Bruyere, O., et al. Incidence of long COVID-19 in people with previous SARS-Cov2 infection: a systematic review and meta-analysis of 120,970

patients. *Intern. Emerg. Med.* **2022**, 10.1007/s11739-022-03164-w, doi:10.1007/s11739-022-03164-w.
6. Baraniuk, C. Covid-19: How Europe is approaching long covid. *bmj* **2022**, *376*.
7. Araf, Y.; Akter, F.; Tang, Y.d.; Fatemi, R.; Parvez, M.S.A.; Zheng, C.; Hossain, M.G. Omicron variant of SARS-CoV-2: genomics, transmissibility, and responses to current COVID-19 vaccines. *Journal of medical virology* **2022**, *94*, 1825-1832.
8. Samani, A.; Mathiassen, S.E.; Madeleine, P. Cluster-based exposure variation analysis. *BMC Med Res Methodol* **2013**, *13*, 54, doi:10.1186/1471-2288-13-54.
9. Canas, L.D.S.; Molteni, E.; Deng, J.; Sudre, C.; Murray, B.; Kerfoot, E.; Antonelli, M.; Rjoob, K.; Pujol, J.C.; Polidori, L. Profiling post-COVID syndrome across different variants of SARS-CoV-2: a prospective longitudinal study on the unvaccinated wild-type, unvaccinated alpha, and vaccinated delta-variant populations. *The Lancet Digital Health* **2023**.
10. Barbagallo, M.; Citarrella, R.; Dominguez, L.; Giannitrapani, L.; Licata, A.; Mansueto, P.; Soresi, M.; Veronese, N. COMEPA (COVID-19 Medicina Policlinico Palermo): a study in hospitalized patients. *Geriatric Care* **2021**, *7*, 9895.
11. Likas, A.; Vlassis, N.; Verbeek, J.J. The global k-means clustering algorithm. *Pattern recognition* **2003**, *36*, 451-461.
12. Milligan, G.W.; Cooper, M.C. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* **1985**, *50*, 159-179.
13. Chen, C.; Haupert, S.R.; Zimmermann, L.; Shi, X.; Fritsche, L.G.; Mukherjee, B. Global Prevalence of Post COVID-19 Condition or Long COVID: A Meta-Analysis and Systematic Review. *The Journal of Infectious Diseases* **2022**.
14. Stavem, K.; Ghanima, W.; Olsen, M.K.; Gilboe, H.M.; Einvik, G. Prevalence and determinants of fatigue after COVID-19 in non-hospitalized subjects: a population-based study. *International Journal of Environmental Research and Public Health* **2021**, *18*, 2030.
15. Natarajan, A.; Shetty, A.; Delanerolle, G.; Zeng, Y.; Zhang, Y.; Raymont, V.; Rathod, S.; Halabi, S.; Elliot, K.; Phiri, P. A systematic review and meta-analysis of Long COVID symptoms. *medRxiv : the preprint server for health sciences* **2022**.
16. Han, Q.; Zheng, B.; Daines, L.; Sheikh, A. Long-Term sequelae of COVID-19: A systematic review and meta-analysis of one-year follow-up studies on post-COVID symptoms. *Pathogens* **2022**, *11*, 269.
17. Blevins, C.A.; Weathers, F.W.; Davis, M.T.; Witte, T.K.; Domino, J.L. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of traumatic stress* **2015**, *28*, 489-498.
18. Bjelland, I.; Dahl, A.A.; Haug, T.T.; Neckelmann, D. The validity of the Hospital Anxiety and Depression Scale: an updated literature review. *Journal of psychosomatic research* **2002**, *52*, 69-77.
19. Shebl, E.; Mirabile, V.S.; Sankari, A.; Burns, B. Respiratory failure. In *StatPearls [Internet]*, StatPearls publishing: 2022.
20. Fan, E.; Brodie, D.; Slutsky, A.S. Acute respiratory distress syndrome: advances in diagnosis and treatment. *Jama* **2018**, *319*, 698-710.
21. Linn, B.S.; Linn, M.W.; Gurel, L. Cumulative illness rating scale. *Journal of the American Geriatrics Society* **1968**, *16*, 622-626, doi:10.1111/j.1532-5415.1968.tb02103.x.

22. Evans, R.A.; McAuley, H.; Harrison, E.M.; Shikotra, A.; Singapuri, A.; Sereno, M.; Elneima, O.; Docherty, A.B.; Lone, N.I.; Leavy, O.C. Physical, cognitive, and mental health impacts of COVID-19 after hospitalisation (PHOSP-COVID): a UK multicentre, prospective cohort study. *The Lancet Respiratory Medicine* **2021**, *9*, 1275-1287.
23. Davis, H.E.; Assaf, G.S.; McCorkell, L.; Wei, H.; Low, R.J.; Re'em, Y.; Redfield, S.; Austin, J.P.; Akrami, A. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine* **2021**, *38*, 101019.
24. West, E.; Moore, K.; Kupeli, N.; Sampson, E.L.; Nair, P.; Aker, N.; Davies, N. Rapid review of decision-making for place of care and death in older people: lessons for COVID-19. *Age and Ageing* **2021**, *50*, 294-306.
25. Sadegi, K.; Soofi, D.; Mirshekarpour, H.; Pormasoumi, H. Investigating the Return Possibility of Covid-19 Mutated Strains and Role of Vaccination in Present and Future. *Journal of Pharmaceutical Negative Results* **2022**, 2043-2047.
26. Pedersen, M.L.; Handberg, C.; Dreyer, P. Mental health reported in adult invasive home mechanical ventilation through a tracheostomy: A scoping review. *International Journal of Nursing Studies Advances* **2022**, 100110.
27. Arulkumaran, N.; Brealey, D.; Howell, D.; Singer, M. Use of non-invasive ventilation for patients with COVID-19: a cause for concern? *The Lancet Respiratory Medicine* **2020**, *8*, e45.
28. Talic, S.; Shah, S.; Wild, H.; Gasevic, D.; Maharaj, A.; Ademi, Z.; Li, X.; Xu, W.; Mesa-Eguiagaray, I.; Rostron, J. Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis. *bmj* **2021**, *375*.
29. Venkatesan, P. Do vaccines protect from long COVID? *The Lancet Respiratory Medicine* **2022**, *10*, e30.