Biopsychosocial Science and Medicine Publish Ahead of Print
DOI:10.1097/PSY.0000000000001380

Unmasking the psychological landscape of Long COVID: a cluster-analytical approach.

AUTHORS: Raquel Gómez Bravo, MD, MSc, PhD[1-2]; Alexandre Infanti, MSc, PhD[2]; Joël Billieux, MSc, PhD[3]; Mark Ritzen, MD [1]; Sandra Nischwitz, MD[4]; Angelika Erhardt, MD[4]; Thérèse Staub, MD[5]; Christian G. Huber, MD, PhD[6-7]; Undine Lang, MD, PhD [6-7]; Steffi Weidt, MD, PhD [8-9-10]; Kira Wolff, MD, MPH, PhD [11]; Juliane Müller, MD, PhD [12]; Claus Vögele, MSc, PhD[2*]; Charles Benoy MSc, PhD[1-6-7*].

*shared last authorship

Corresponding author:

Dr. phil. Charles Benoy
Email: charles.benoy@chnp.lu
17, avenue des Alliés
B.P. 111 · L-9002 Ettelbruck

AFFILIATIONS:

1 Centre Hospitalier Neuropsychiatrique (CHNP), Rehaklinik, Ettelbruck, Luxembourg
2 Department of Behavioural and Cognitive Sciences, University of Luxembourg, Esch-sur-Alzette, Luxembourg
3 Institute of Psychology, University of Lausanne (UNIL), Lausanne, Switzerland
4 Max-Planck-Institute for Psychiatry, Munich, Germany
5 Department of Infectious Diseases, Centre Hospitalier du Luxembourg
6 Universitäre Psychiatrische Kliniken Basel, Switzerland
7 University of Basel, Switzerland
8 Psychiatry of St. Gallen, Switzerland
9 University of Zurich, Switzerland
10 University of St Gallen
11 Calm Tagesklinik München, Germany
12 Universitätsklinikum Frankfurt am Main, Germany

Source of funding

This project was supported by the Luxembourg National Research Fund (COVID-19/2021/16874342/Psy-Long-COVID).

Conflicts of Interest

None declared.

1 figure, 2 tables, 2 Supplemental Digital Content files (.pdf, .docx)

Article Editor: Daryl O'Connor

ABSTRACT

OBJECTIVE: Since the emergence of SARS-CoV-2, an increasing number of people report long-term physical and psychological impairments. Research on the immunological sequalae of long COVID (LC) is growing, though its relationship with mental health remains underexplored. We investigated the psychological impairments associated with LC, identify related psychological symptom clusters, and their relationship with physical symptoms and pandemic-related variables.

METHODS: Cross-sectional descriptive study, using an online questionnaire (September 2020 - December 2022) in German and French, assessing depression, anxiety, fatigue, stress, and somatic symptoms. Clusters were identified using hierarchical and machine learning techniques (Kmeans and AffinityPropagation) and compared based on LC symptoms, past physical and mental health, substance use, COVID-19 variant, and family dynamics.

RESULTS: Among 1218 LC participants (78.7% female), four clusters were identified using AffinityPropagation: (1) low anxiety, depression, and somatoform symptoms, (2) low anxiety and depression but moderate somatoform symptoms, (3) high anxiety and depression with high somatoform symptoms, and (4) high anxiety and depression with moderate somatoform symptoms. Cluster 3 reported the most severe physical and neurological

symptoms, the largest life impact (including relationship deterioration and professional difficulties) and the highest prevalence of past mental disorders (depression and PTSD). Cluster 1 reported the least symptoms.
CONCLUSIONS: LC significantly impacts mental health, particularly through increased anxiety, depression, and somatoform symptoms, especially in patients with severe LC physical symptoms. Psychological interventions targeting distinct symptom clusters may improve both mental and physical outcomes. Early mental health screening and tailored interventions are recommended for LC assessment and treatment.

Keywords: Long COVID (LC); Post Covid Syndrome PCC; Post-acute sequelae of SARS-CoV-2 infection (PASC); Psychological impact; Psychological distress; Mental health; Psychological symptoms; Cluster analysis; AffinityPropagation

INTRODUCTION

Long COVID (LC), also known as post-COVID-19 syndrome or post-acute sequelae of SARS-CoV-2 infection (PASC), is a complex and evolving condition characterized by a wide range of physical and psychological symptoms. These symptoms persist for more than 12 weeks beyond the acute phase of the initial COVID-19 infection and cannot be explained by an alternative diagnosis (1). Long COVID can lead to long-term disability and functional impairment (2), significantly impacting the quality of life of individuals, including children and adolescents (3).

The estimated prevalence of LC varies widely but it is currently estimated to affect from 10% up to 80% of patients who had COVID-19, depending on the population studied and the criteria used (4-6). The degree of disability in LC patients is affected by several factors, including the severity of the initial COVID-19 infection, individual responses to the virus, specific symptoms experienced, gender, social factors, pre-existing mental health conditions, comorbidities, and elevated inflammatory markers (7). Psychological, physical, and cognitive symptoms can persist for weeks or months after the acute phase of the illness. Some individuals may experience mild or intermittent symptoms, while others face more significant disability. The degree of disability can change over time, with some individuals experiencing gradual improvement with appropriate medical care and rehabilitation, while others may experience chronic symptoms, leading to long-term disability (8).

Long COVID can manifest as a variety of physical symptoms affecting multiple systems, accompanied by persistent psychological issues, and neuropsychological impairments. Fatigue and sleep disturbances are commonly reported, along with symptoms like anxiety, depression, stress, post-traumatic stress, headaches, cognitive impairment ("brain fog"), decreased concentration, muscle pain, sensory and motor deficits, among others (9-11). Similar symptoms have been observed in various viral infections, affecting both physical and psychological aspects, as well as presenting neuropsychological challenges (12). These symptoms can have a significant impact on an individual's quality of life, ability to work or attend school, and mental well-being (13). They are collectively referred to as chronic post-viral syndromes, describing a cluster of persistent symptoms and complications that individuals often experience following recovery from a viral infection (14).

Recent research has emphasized that fatigue is the most prominent physical symptom affecting daily functioning in LC patients, with 87% of them experiencing it and severely limiting their ability to return to work or seek new employment (15). Nevertheless, LC presents a spectrum of symptoms, both physical and psychological, and cluster analysis has been employed to identify distinct groups of LC patients based on their physical symptom profiles (16-18)

These clusters are associated with different levels of functional impairment and highlight the intricate relationship between physical and psychological symptoms in LC (19). Additionally, research has identified profiles associating long-term post-COVID symptoms, COVID-19 symptoms at hospital admission, and prior medical comorbidities, emphasizing the importance of considering the temporal progression of symptoms in understanding LC (17, 20).

Previous studies identified distinct physical symptom clusters, such as gastrointestinal, musculoskeletal, neurocognitive, airway, and cardiopulmonary, and there is evidence for associations between these clusters and an increased risk of depression and anxiety, highlighting the psychological consequences of LC (21 - 24). Some findings suggest that LC may be associated with comorbid mental disorders, such as anxiety and depression, with higher rates in those with persistent symptoms (25,26). Nevertheless, only few studies have explored psychiatric sequelae in LC patients, or the factors associated with them (27).

People with debilitating and disabling medical conditions (i.e., cancer and asthma) tend to experience higher rates of mental comorbidity (28) compared to the general population. As pointed out by Re'em et al. (27), LC may also be associated with comorbid mental health problems. Previous studies have identified elevated levels of anxiety and depression symptoms in those who have had COVID-19, with milder cases exhibiting lower rates in contrast to more severe cases (28).

In summary, LC patients report a wide range of psychological symptoms that may arise from various factors, such as the direct effects of the immune response to the initial infection, and the realization of having a condition that persists over an extended period of time (as in chronic disorders), or an interaction of both. Nonetheless, there is a risk of misdiagnosis when assuming all symptoms are solely due to psychiatric issues (29-31), which can contribute to increased stigmatization (32).

This study investigates the relationship between psychological distress and physical symptoms in Long COVID (LC) patients, aiming to identify distinct subgroups based on their psychological profiles, as assessed with standardized scales. While previous studies have often focused on isolated psychological symptoms for analysis, the interrelation between these symptoms and their association with physical manifestations of LC remains largely unexplored. Thus, this study seeks to determine how these symptom clusters differ not only in terms of their psychological symptom profiles but also in physical symptoms and variables associated with the current pandemic. The identification of such symptom clusters could significantly improve our understanding of the heterogeneity of LC and would be crucial for providing more accurate assessments and comprehensive, effective care that addresses both psychological and physical aspects.

Objectives

1) To identify LC related potential psychological syndromes and categorize individuals into distinct groups based on their psychological symptoms.
2) To explore potential differences in physical symptoms between the identified psychological syndrome clusters.
3) To examine whether syndrome clusters differ in relation to other pandemic-related factors, encompassing psychosocial and medical aspects.

METHODS

Study Design and Setting

We used an online, descriptive, cross-sectional survey with 184 items (Supplement al Digital Content 1, http://links.lww.com/PSYMED/B78). Participation was voluntary, anonymous, and participants could withdraw at any point. The self-report questionnaire was made available in both German and French, but not confined to these countries. Completing the questionnaire typically took an average of 30 minutes. We selected respondents based on the World Health Organization's definition of LC, which encompasses individuals who have newly developed symptoms three months after their initial SARS-CoV-2 infection, with these symptoms persisting for a minimum of two months without any other identifiable explanation (33).

The survey was online from February 2020 to December 2022. It included questions on participants' global LC syndrome, considering both current and past symptoms, with an emphasis on their severity, duration, and impact (loss of taste/smell, other ear-/nose and throat disorder, headache, migraines, memory or concentration, other neurological symptoms, cardiorespiratory, gastrointestinal symptoms, vascular/lymph node symptoms, skin symptoms and others). The non-standardized items were derived from a thorough review of the literature on LC and refined through consultations with clinical experts to ensure they captured the most frequently reported symptoms associated with LC. Although formal content validation was not conducted, these items were reviewed for clarity and relevance by bilingual experts, ensuring both cultural and linguistic appropriateness in German and French. In addition, the survey assessed mental health, the impact of health on life quality, perceived stress, traumatic experiences, and somatic symptom experiences. In detail, this included an assessment of depression (PHQ-4, i.e. 4 items of the PHQ-9 without physical symptoms), anxiety (GAD-7), fatigue (FAS), psychological experiences related to physical symptoms (SSEQ) and perceived stress (PSS). Additionally, we explored factors relating to individuals and the pandemic, including sociodemographic factors, vaccination status, time of infection, the acute phase, acute treatment setting, health-related factors, and the impact of the pandemic on various aspects of personal life (CEFIS scale). To account for potential pre-existing conditions, the survey included a comprehensive assessment of potentially confounding variables. Furthermore, we collected data on a range of possible risk factors that may have occurred chronologically prior to the Coronavirus infection (see 34, for the entire study protocol).

Sampling and Recruitment

We employed snowball sampling as our recruitment method, encompassing public relations and information campaigns aimed at both the general public and healthcare professionals. Additionally, we established collaborative partnerships with various specialized units, including LC units, intensive care units, infectious diseases departments, and mental health care departments such as psychiatry units. This approach facilitated the direct referral of individuals with LC to our research project. This study is part of a larger, multi-center, and international initiative, and the recruitment procedure remained consistent across all participating institutions.

Ethical Approval:

The study adhered to the principles outlined in the Declaration of Helsinki. Ethical approval was not deemed necessary since the data collected was non-identifiable and anonymous. This approach was confirmed by the relevant boards of the participating institutions, such as the Ethical Commission of the Ludwig-Maximilian-University (Munich, Germany), the Ethics Committee for Northwest and Central Switzerland (EKNZ), and the Ministry of Health in Luxembourg.

Virus Variants:

The assignment of specific SARS-CoV-2 variants to participants was based on information on the reported country of residence and the reported infection date. To estimate the virus variant prevalent at the reported date of infection, we used data from the GISAID initiative (35). The specific data source for this estimation was the GISAID variant distribution tracker, which offers a comprehensive, up-to-date overview of the variant distribution across different countries. This approach allowed for a more accurate approximation of the likely virus variant each participant had contracted, considering the temporal and geographical prevalence of specific variants as per the GISAID database.

Statistical analysis

Standard data cleaning procedures were applied before analysis. This included reviewing the data for missing or incomplete responses, removing any duplicates, and ensuring the validity of the responses. Only completed surveys were included in the analysis. This study employed two distinct clustering approaches using the total scores of the GAD7, PHQ-4 (PHQ-9 excluding items which refer to physical symptoms), and all the SEEQ dimensions (34). The first approach involved a combination of hierarchical and non-hierarchical clustering methods, using NbClust and Kmeans clustering to compensate for the weakness of each other, following the recommendations provided by Hair and colleagues (36). In this procedure, we first applied a hierarchical cluster to obtain a suggested number of clusters to be used in the further Kmeans clustering. For this purpose, we used a R package (v4.2.0), named NbClust (37). After this step, we proceeded to a Kmeans clustering analysis to generate a suggested number of clusters identified in the previous step.

The second approach employed the AffinityPropagation machine learning model with parameter tuning for cluster generation. In this approach, we tuned a parameter named "preference", which influences the number of clusters generated. After generating the clusters with both approaches, we compared the two resulting cluster solutions using three intrinsic indexes (silhouette, Calinski-Harabasz, Davies-Bouldin). In addition, we assessed the cluster solutions from a clinical perspective.

Once the final cluster solution was chosen, we compared the identified clusters on the following external variables: LC symptoms, previous medical illness, previous mental health problems, life impact, substance use, vaccination status, COVID-19 variant, stress level, pain and suffering levels, gender, and family situation. The covariates used in the modelling were standardized prior to analysis. This step was taken to ensure that variables with different ranges, such as symptom scores, did not disproportionately influence the clustering process. By standardizing the covariates, we minimized the risk of symptoms with wider scoring ranges dominating the analysis and ensured a more balanced contribution of each variable to the clustering model.

Cluster Comparisons and Statistical Analysis:

To compare clusters with respect to physical symptoms, we generated contingency tables and performed Chi-Squared tests complemented by Cramer's V analysis to report effect sizes. Post hoc analyses were conducted using the chisq.post.hoc.test R package (38), applying Bonferroni correction based on residuals from Pearson's Chi-squared Test for Count Data.

This procedure was also applied to explore differences between syndrome clusters in terms of pandemic-related variables, including psychosocial and medical factors. For continuous external variables, we used the Kruskal–Wallis test followed by Dunn's test as a post hoc analysis with Bonferroni correction.

Analysis of Symptom Distributions Across Clusters:
To examine the symptom distributions across the four clusters, we conducted a Chi-Squared test with 6 degrees of freedom. Symptoms were categorized into three distinct groups (e.g., mild, moderate, severe). While these categories may suggest an ordinal structure, they were treated as unordered in the analysis to allow for a more flexible evaluation of differences between clusters. This approach avoids imposing assumptions about linearity or equal spacing between categories, focusing instead on capturing broad patterns and disparities in symptom distributions.

The primary goal of this analysis was to identify and interpret meaningful differences across clusters, including potential non-linear or complex relationships. Treating the symptom categories as unordered ensures that all patterns in the data are considered, preventing oversimplifications that may arise from tests for trends assuming monotonic changes.

For a more detailed description of the study protocol, methodology and data-analytic strategy, see Gómez Bravo et al. (34).

RESULTS

Participants included 1218 individuals with self-identified LC. The mean age was 42.61 years old (SD=11.82) with a majority being female (78.7%, N=959). Participants mostly came from Germany (77.1%, N=939) and Luxembourg (15.8%, N=192). A majority of the sample were professionals (75.6%, N=920), refer to individuals engaged in skilled occupations, and had a university degree (56.9%, N=693) or vocational training (27.3%, N=332).

The Kmeans and AffinityPropagation clustering solutions are both reported in Fig 1. While the Kmeans solution proposed a three-cluster solution, the AffinityPropagation solution proposed a four-cluster solution. Regarding the indexes, the AffinityPropation presented better silhouette (0.36 vs 0.35) and Davies-Bouldin (1.42 vs 1.46) scores, but a lower Calinski-Harabasz score (424.15 vs 513.12).

- Figure 1 here -

While the Kmeans solution identified only one cluster characterized by a high level of anxiety and depression scores, the AffinityPropagation solution proposed two that differ in their psychological experiences related to physical symptoms. From a clinical perspective, this distinction is important since it implies that the level of perceived burden imposed by physical symptoms is not necessarily associated with increased levels of anxiety and depression. Moreover, the four clusters solution proposed by the AffinityPropagation model was also supported by the intrinsic indexes. We thus decided to use the four-cluster solution, based on both psychological (anxiety and depression) and physical symptom severity (e.g., somatoform symptoms), as well as on clinical judgment, as this offers a more nuanced understanding of psychological and somatic symptom experiences, prioritizing clinical relevance over statistical criteria. The following profiles were proposed:

a) Low levels of anxiety and depression symptoms and low level of somatoform symptoms (cluster 1).

b) Low levels of anxiety and depression symptoms and moderate levels of somatoform symptoms (cluster 2).

c) High levels of anxiety and depression symptoms and high levels of somatoform symptoms (cluster 3).

d) High level of anxiety and depression symptoms and moderate levels of somatoform symptoms (cluster 4).

Participants with severe fatigue (upper quartile FAS scores) were 3.2 and 2.8 times more likely to meet clinical thresholds for anxiety (GAD-7) and depression (PHQ-4), respectively, compared to participants with milder fatigue symptoms. This relationship was particularly pronounced in Cluster 3, which exhibited the most severe physical symptoms alongside the highest prevalence of anxiety and depressive symptoms.

Long Covid symptoms

There were no gender differences between cluster. Clusters differed on all LC symptoms ($p < .001$), except for two: loss of taste/smell and Skin symptoms. Clusters 2 and 3 presented the highest rate of headache/migraines symptoms, while cluster 1 showed the lowest LC symptoms when compared with other clusters. Participants in cluster 3 reported more neurological symptoms and vascular/lymph node symptoms than those in the other clusters (Table 1).

- Table 1 here -

These clusters differed in terms of physical LC related symptoms and medical and mental health history. Related to previous medical illness, the results suggest that while some medical conditions are evenly distributed across the clusters, others vary significantly, indicating potential differences in the medical backgrounds of the clusters (Supplemental Digital Content 2, http://links.lww.com/PSYMED/B79). This could be important for understanding the heterogeneity of the long COVID patient population and tailoring medical care to each group's specific needs.

In contrast, the analysis showed a significant difference in the prevalence of epilepsy between clusters ($p = .023$). Cluster 3 had the highest proportion of individuals with epilepsy at 2.4%, while cluster 2 had no reported cases. Clusters 1 and 4 had proportions of 0.9% and 0.4%, respectively.

The category “others”, which includes illnesses not specified in the list, showed a significant difference ($p = .040$). Cluster 3 had the highest proportion at 34.1%, while cluster 1 had the lowest at 24.8%.

For the other conditions listed (such as diabetes, hypertension, obesity, cardiovascular diseases, respiratory and lung conditions, asthma, kidney disease, liver disease, cerebrovascular disease, cancer, organ transplant, allergies), the differences were not statistically significant, as indicated by higher p-values (above .05). For these conditions, the distribution across the four clusters did not differ more than would be expected by chance.

Significant differences were also identified between the four clusters in terms of previous mental health problems. Past depression, for example, was reported more frequently in clusters 3 and 4 (46.5% and 41.2%, respectively) compared to clusters 1 and 2 (19.1% and 21.6%, respectively) (chi-square $p < .001$; Cramer's V.245, suggesting a moderate effect size).

Similarly, previous post-traumatic stress disorder shows a higher prevalence in cluster 3 (20%) compared to the other clusters, with the chi-square test also indicating a significant difference ($p < .001$) and a Cramer's V of .151, suggesting a moderate effect size.

There are further differences between clusters in terms of previous mental health problems:

- ADHD: Cluster 4 showed the highest prevalence of ADHD at 5.3%, which was significantly higher than the other clusters ($p = .007$).
- Burnout: Cluster 3 had a notably higher prevalence of burnout at 17.1%, which was significantly different from clusters 1 and 4 ($p = .005$).
- Anxiety Disorders: Cluster 3 showed the highest prevalence at 34.1%, which is markedly higher than Cluster 1 which had 9% ($p < .001$).
- Adjustment Disorders: Again, Cluster 3 exhibited a higher prevalence at 10%, which was significantly different from clusters 1 and 4 ($p = .005$).
- Phobia: Cluster 3 showed a higher prevalence of phobia at 6.5%, which was significantly different from Cluster 1 with 1.4% ($p < .001$).
- Obsessive-Compulsive Disorder (OCD): Cluster 3 had a higher prevalence of OCD at 4.7%, which was significantly different from Cluster 2 which had 0.8% ($p = .021$).

These results suggest that participants in Cluster 3 were significantly more affected by mental disorders in the past.

Life impact

In terms of life impact, clusters significantly differ in their response to whether they could live in their current state of health in the long term (Supplemental Digital Content 2, http://links.lww.com/PSYMED/B79). Cluster 3 has the lowest percentage of individuals who feel they could live with their current health state in the long term (6.5%), significantly less than Cluster 1 (38.1%) and Cluster 4 (20.4%).

Regarding maintaining the same life rhythm as before, Cluster 2 has the highest percentage of individuals who answered 'no' at 80.2%, indicating a significant impact on their lifestyle compared to Cluster 1 (62.6%).

When it comes to returning to professional activity without complications, Clusters 2 and 3 show higher percentages of individuals who felt that they could not return to professional activity (55.4% and 54.1%, respectively) compared to Clusters 1 and 4 (42.1% and 42.9%, respectively).

Changes in relationships that have deteriorated are most significant in Clusters 3 and 4 (74.1% and 64.2%, respectively), indicating a more negative impact on personal relationships compared to Cluster 1 (35.2%) and Cluster 2 (51.9%).

In conclusion, regarding life impact, those in Cluster 3 are the most negatively affected by LC in that they report the lowest capacity to live in their current state in the long term, the most dramatic changes in life rhythm and the most negative effects on relationships. Those in cluster 1 were least negatively affected by LC (Supplemental Digital Content 2, http://links.lww.com/PSYMED/B79).

Substance abuse

Significant differences regarding substance use were identified between clusters. Specifically, the use of medication, nicotine, psychotropic drugs, and alcohol shows variation across clusters. For example, the use of medication is highest in Clusters 2 and 3, with 65.4% and 66.5% respectively, and lower in Clusters 1 (50.4%) and Cluster 4 (54%). The use of psychotropic drugs is higher in Cluster 4 (3.5%) compared to other clusters and the use of nicotine is higher in Clusters 3 and 4 (15.9% and 15.5%, respectively) compared to Clusters 1 and 2 (10.4% and 9.3%, respectively). Alcohol abuse is also higher in Cluster 4 (7.1%) and Cluster 3 (5.3%) compared to Clusters 1 and 2 (Supplemental Digital Content 2, http://links.lww.com/PSYMED/B79).

Vaccination status

While the percentage of individuals who reported not to be vaccinated was highest in Cluster 3 (13.5%), the difference to the other clusters was only marginal (p = .067), suggesting that vaccination status did not significantly differ between clusters (Supplemental Digital Content 2, http://links.lww.com/PSYMED/B79).

Family situation

A significant difference between clusters occurred with regards to the variable family situation, although there was no clear pattern. This concerned mainly the percentage of individuals who reported to be single with children, which differed between the clusters, with Cluster 3 having the highest (9.4%) and Cluster 1 the lowest percentage (5.4%). There are also differences in the percentages of individuals who are single without children, although the proportions are relatively similar across the clusters.

The distribution of individuals living as a couple with children varies significantly, with Cluster 1 having the highest percentage (45.4%) and Cluster 4 the lowest (34.5%). There are further differences, with Cluster 4 reporting the highest percentage of couples without children (30.5%) and Cluster 2 having the lowest (25.3%).

The percentage of individuals living with their parents also differs, with Cluster 3 showing the highest percentage (9.4%) and Cluster 2 the lowest (3.5%). Finally, differences are observed in shared housing situations, with Cluster 4 having a higher percentage (4.9%) compared to Cluster 1 (1.7%). The chi-square test indicates that these differences are statistically significant (p = .034). Cramer's V (.085) suggests a low effect size for cluster differences concerning family situation (Supplemental Digital Content 2, http://links.lww.com/PSYMED/B79).

Table 2 summarizes the characteristics with which the respective cluster profiles of psychological syndromes could be associated with. In the analysis of Long COVID symptom severity across clusters, the symptom questions were categorized into three ordinal levels, representing increasing severity. Initially, these categories were analyzed using contingency tables, with degrees of freedom calculated as (rows−1)×(columns−1)=6(rows - 1) \times (columns - 1) = 6(rows−1)×(columns−1)=6. To better capture the trends inherent in the data, we conducted trend analyses using the

Cochran-Armitage test for unadjusted comparisons and ordinal logistic regression to model the associations between symptom severity and cluster membership. This revised analysis revealed significant trends, with participants in Cluster 3, characterized by the most severe physical and psychological symptoms, being significantly more likely to report higher severity levels for symptoms such as fatigue and neurological impairments compared to other clusters.

- Table 2 here -

DISCUSSION

The present study was designed to investigate the relationship between psychological distress and physical symptoms in Long COVID (LC) patients, aiming to identify distinct subgroups based on their psychological profiles, as assessed with standardized scales. Previous research has shown a positive relationship between psychosocial distress and physical symptoms (39) in survivors of Ebola virus disease. For LC, an analogous relationship has been suggested and our study expands this understanding specifically in this context. A better understanding of such symptom clusters could significantly improve assessment and, therefore, comprehensive, effective care that addresses both psychological and physical aspects of those with LC.

Previous research has employed cluster analysis to identify distinct groups of LC patients based on their symptom profiles. For instance, Kenny et al. (19) described distinct clinical phenotypes or clusters of symptoms in LC patients, associated with functional impairments, including longer work absence, higher dyspnea scores, and lower scores in domains of general health, physical functioning, and social functioning. This suggests that there may be different underlying pathophysiological mechanisms driving these symptom clusters, underscoring the intricate relationship between physical and psychological symptoms in LC. There is evidence that LC patients experience persistent depression and anxiety symptoms (21). This underlines the need for monitoring mental health and providing adequate support alongside the diagnosis and treatment of the physical consequences of LC. Fernández-de-las-Peñas et al. (22) focused on previously hospitalized COVID-19 survivors and used cluster analysis to identify profiles associating long-term post-COVID symptoms, COVID-19 symptoms at hospital admission, and prior medical co-morbidities, highlighting the importance of the temporal progression of symptoms in understanding LC. Goldhaber et al. (23) conducted an analysis of physical symptom clusters and risk factors for post-acute sequelae of COVID-19. They identified five symptom clusters (gastrointestinal, musculoskeletal, neurocognitive, airway, and cardiopulmonary) and found associations between certain symptom clusters and increased risk of depression and anxiety, emphasizing the psychological manifestations of LC. Tsuchida et al. (24) also classified five clusters based on the symptomatology, in a cross-sectional study in Japan. Notably, clusters in which fatigue emerged as a prominent symptom displayed a correlation with the severity of other symptoms, implying the potential influence of social factors on the manifestation of LC. Although these studies provide preliminary evidence for mental ill-health in those with LC, none of these studies used standardized questionnaires to assess psychological syndromes. It remains largely unclear, therefore, to what extent the identified clusters are reliable in terms of assessment and clinical relevance. An exception is Re’em et al. (27) who recently published a study addressing psychiatric symptomatology assessed using standardized questionnaires to phenotype LC patients.

The present study used a similar approach, although we carried out an online study, without control group, performing both, a Kmeans and AffinityPropagation clustering analysis, which - to the best of our knowledge - has not been carried out before using questionnaire based self-reports. This dual-method approach, especially the use of AffinityPropagation, is innovative as it likely provides a more nuanced understanding of psychological symptom clustering in LC patients. Re'em et al. (27) focused more on psychiatric outcomes and their associations with non-psychiatric symptoms and social factors, whereas this study emphasizes the clustering analysis of online self-report questionnaires. This approach allows for a broader understanding of psychological distress in LC, encompassing both psychiatric and subclinical psychological symptoms. Moreover, our study identified four distinct clusters of LC patients with different levels of anxiety, depression, and somatoform symptoms. This fine-grained differentiation provides a more detailed understanding of the psychological landscape of LC patients, offering new insights into the complex interplay of different psychological symptoms in LC, which can inform more personalized treatment approaches.

We opted for a four-cluster solution as suggested by the AffinityPropagation model and supported by the intrinsic indexes and a clinical assessment. The model suggested two clusters groups which show either high or low levels of anxiety and depression symptoms. Between these two groups, clusters can be differentiated in terms of their levels of somatoform symptoms. In this sense, the novelty of our contribution also lies in integrating physical symptoms and medical history with psychological profiles. This comprehensive approach provides a more comprehensive view of the patient, emphasizing the interconnected nature of physical and mental health in LC. This integration is more extensive and detailed compared to Re'em et al. (27), highlighting the importance of considering both physical and psychological symptoms in understanding and managing LC, with implications for personalized healthcare strategies in LC. By understanding the unique attributes of each cluster, healthcare professionals can tailor treatments and support measures more effectively. These clusters differed in terms of physical LC related symptoms and medical and mental health history. Previous medical health problems significantly differ between clusters, as do numerous previous mental disorders.

This severity-based clustering aligns with other analyses in the Long COVID research literature (40). Our findings also support the view that the intensity of symptoms—both psychological and physical—often co-varies, which can help guide tailored treatment approaches for different severity levels. In fact, the clusters with strong physical symptomatology are related to a higher prevalence of previous mental health problems. Specifically, Cluster 3, which exhibits the most severe physical symptoms, also shows the highest prevalence rates for several mental health conditions, including depression, post-traumatic stress syndrome, and anxiety disorders. This indicates a strong correlation between severe physical symptoms and a history of mental health problems. Cluster 4, while similar to Cluster 2 in having moderate to high rates of certain mental health conditions, also shows a notable correlation between physical symptoms and previous mental health problems, although this relationship is less intense compared to Cluster 3. In summary, the more severe the physical symptomatology in a cluster, the higher the likelihood of a significant history of mental health issues. This relationship underscores the interconnected nature of physical and mental health, particularly in the context of LC. Fatigue, a hallmark symptom of LC, not only diminishes physical functioning but also significantly correlates with higher levels of anxiety and depression, likely driven by its pervasive impact on patients' ability to maintain daily activities and long-term goals, suggesting that fatigue not only reflects physical illness severity but may also exacerbate psychological distress by impairing daily functioning and increasing the perception of long-term disability. This underscores the importance of integrated care approaches that address both physical and mental health in the management of Long COVID.

Clusters also significantly differ in terms of life impact, substance use, and COVID-19 variant. Participants allocated to cluster 3 reported more symptoms, and this was associated with more frequent reported illnesses and a lower capacity to cope in the long term. There are no significant differences in vaccination status or gender among clusters. Lastly, clusters vary in terms of family situations. These findings suggest that the clusters with more severe physical symptomatology (especially Cluster 3) also tend to have a greater impact on life, higher substance use, and distinct relationships with COVID-19 variants. This result, however, should be interpreted with caution as we only inferred viral strain from the self-reported time of infection (Supplemental Digital Content 2, http://links.lww.com/PSYMED/B79).

Several studies on LC treated psychological symptoms uniformly, without much differentiation of the types of mental distress experienced. Our findings echo previous research, showing a correlation between psychological symptom severity and the severity of physical LC symptoms. The current results further corroborate previous findings, showing a significant relationship between severe physical symptoms and prior mental health issues. Less severe psychological symptoms tend to accompany milder Long Covid physical manifestations (Clusters 3 and 1).

Our research also provides new insights. It appears that LC can lead to diverse psychological syndromes not directly tied to the severity of physical symptoms. Clusters 2 and 4 demonstrate this variety, with Cluster 2 leaning towards somatoform disturbances and Cluster 4 more towards anxiety and depression. These differences are associated with factors beyond the experience of physical symptoms. For instance, Cluster 2 individuals often face greater challenges in functional and professional aspects, while Cluster 4 individuals experience more pronounced interpersonal and social difficulties.

Although the link between somatoform symptoms with functional and occupational issues, and between interpersonal-social and anxious-depressive symptoms seems psychologically intuitive, our findings suggest for the first time that the mental distress associated with LC is not just a reflection of the severity of the physical illness. It is also related to its perceived impact on various levels of functionality and aspects of individuals' lives.

Physical symptoms of Long COVID (LC) patients fluctuate normally during the day and can interact with psychological symptoms like anxiety and depression. This daily variability in physical health can exacerbate psychological distress, creating a bidirectional relationship between them. Additionally, the stability of the identified psychological clusters over time may be impacted by these fluctuations, as patients could shift between clusters based on changes in their physical symptoms. Future longitudinal studies are needed to assess the stability of these clusters and the need for personalized, adaptive treatment approaches that address both physical and psychological symptoms.

Limitations

It is important to note that the study may have attracted mainly participants who were personally affected by impairments, contributing to self-selection bias as described in the protocol (34). Given our recruitment strategy, which focused on Long COVID units and mental health care departments and employed snowball sampling, the sample is likely biased toward individuals with more severe symptoms or those already engaged in care. This approach limits the generalizability of our findings to the broader Long COVID population, particularly those who may not seek specialized care or experience milder symptoms. Consequently, the extent to which these results are applicable to all individuals with Long COVID cannot be determined. Due to the snowball sampling method used in this study, it was not possible to determine how many potential respondents were approached, making it difficult to calculate a precise response or completion rate. This poses a limitation in terms of the generalizability of the results, as the sample may not fully represent the broader Long COVID population. The duration of the study period might have also impacted the results, as the common understanding of the condition has been evolving since the beginning of the pandemic. Additionally, the online format may have made it difficult to control and observe participants during its completion. Moreover, this study presents the same bias as any online research such as accuracy of the assessment, accessibility barriers, digital divide, limited interactivity, and participant selection bias etc.

The lack of official confirmation of infections, symptoms, and impairments before and after the SARS-CoV-2 infection makes it challenging to compare differences accurately, relying solely on participants' reports, making accurate comparisons challenging. Moreover, there was a gender bias, as the majority of participants were female, which could affect the results due to higher prevalence rates of mental health problems in women and post-viral illness. Another limitation of the study is that no formal content validation was conducted. The questionnaire was developed during the pandemic by clinical experts who reviewed the items and reached a consensus that the list of symptoms was appropriate. The experts assessed whether the items captured the key physical and psychological symptoms of Long COVID based on the available knowledge, ensuring clarity and accessibility for participants. The bilingual clinical experts ensured accuracy and cultural relevance in the German and French versions, maintaining consistency with the original meaning. The rest of the consortium, consisting of native speakers, confirmed the accuracy between the two versions, allowing for reliable comparisons across language groups. The items were meant to supplement standardized scales used in the study, such as PHQ-4 and GAD-7, rather than serve as standalone diagnostic tools. The lack of physiological measurements and information on the time since infection and associated virus variants further limit the study's conclusions. Another important limitation is the lack of a control group. Finally, it is essential to note that the study did not conduct follow-up assessments to determine symptom duration or possible improvements or deteriorations.

Further Research

Comparing subjective impairments with objective measurements could provide further understanding of how individuals assess their abilities. Additionally, the prevalence of different virus variants, vaccination status, previous medical conditions and treatment of the participants should be considered to explore their effects on cognitive and psychological impairments. Further longitudinal research and larger cohorts should be considered to analyse the long-term consequences of LC on cognitive functions and mental health, including psychological and psychiatric interventions, ensuring continuity of care in those experiencing various psychological syndromes.

CONCLUSIONS

The findings of this study highlight the significant impact of LC on individuals' mental health, and the related adverse consequences for quality of life, and productivity. However, the clusters described in this study demonstrated substantial impact on quality of life and productivity associated with this condition and underscore the urgency to improve diagnosis and treatment without further delay.

A conclusion to be drawn from these results is that early screening, treatment, and support for different groups of psychological symptoms may improve the overall condition of LC patients. Psychological interventions can improve mental well-being, which, in turn, has a beneficial effect on physical health outcomes. Managing psychological distress is crucial as it has the potential to exacerbate physical symptoms or impede the recovery process. Addressing these psychological factors may well alleviate some of the physical manifestations associated with LC. Furthermore, it could enhance patients' ability to cope with their physical symptoms, adhere to treatment regimens, and actively participate in rehabilitation, which would contribute to an improvement in overall symptoms. Our results also suggest that psychological ill-health in the context of physical illness should not be considered as a single homogenous entity. Psychological symptoms differ significantly between individuals affected by the same infectious disorder, which needs to be taken into account in clinical practice. Finally, providing early psychological support might also help in reducing the severity and progression of LC symptoms over time, thereby potentially yielding better long-term health outcomes.

Another conclusion from the present results concerns the use of online testing in research and clinical practice. Online psychological testing can play an important role in assessing the psychological impact of a wide range of disorders. It can help identify and assess mental health symptoms, cognitive impairments, and provide accessible support for individuals experiencing persistent symptoms. Individuals can access psychological testing and support from their own homes, making it a convenient and effective tool for addressing the mental health needs of those affected and providing targeted interventions accordingly.

Author contributions

CB and CV designed the study. CB and RGB conducted the literature review and led the data collection. AI carried out the statistical analysis and clustering modelling, advised by JB. RGB and AI drafted the manuscript and CB, CV and JB edited it prior to submission, making substantial contributions to it. MR, SM, SN, AE, TS, CH, UL, SW, KW and JM were involved in revisions of the study design and actively participated in the recruitment of participants. All authors reviewed the manuscript and approved the final version.

Acknowledgements

We gratefully acknowledge permission to freely use the scales FAS, PSS, CEFIS, PHQ, SSEQ and GAD.

We extend our gratitude to the following partners for their contributions and active involvement in this study: Zentrum fir Psychotherapie, Rehaklinik, Centre Hospitalier Neuro-Psychiatrie (CHNP), Ettelbruck (Luxembourg), University of Luxembourg (Luxembourg), University of Lausanne (UNIL), Centre Hospitalier du Luxembourg (CHL), Max-Planck-Institut für Psychiatrie, Munich (Germany), Psychiatrische Universitätsklinik, Zurich (Switzerland), Universitäre Psychiatrische Kliniken, Basel (Switzerland), University of Basel, (Switzerland), Universitätsklinikum Frankfurt am Main (Germany), Universitätsklinikum Hamburg-Eppendorf (Germany).

We gratefully acknowledge our partnership with the national CoVaLux programme funded by the Luxembourg Government. The commitment of this programme to fostering multidisciplinary research has been instrumental in uniting various research initiatives within the same domain and enhancing the scope and impact of our work.

Finally, we extend our heartfelt gratitude to Prof. Moritz for his invaluable collaboration in designing our study and his unwavering dedication to multidisciplinary Long COVID research. His significant contributions have been fundamental in enabling us to conduct exceptional research within the consortium.

Abbreviations

ADHD: Attention-deficit/hyperactivity disorder
CEFIS: COVID-19 Exposure and Family Impact Scales
JMIR: Journal of Medical Internet Research
LC: Long COVID
OCD: Obsessive-compulsive disorder
PTSD: Post-traumatic stress disorder
RCT: randomized controlled trial

REFERENCES

1. Parkin, A., Davison, J., Tarrant, R., Ross, D., Halpin, S., Simms, A., et al. (2021). A multidisciplinary NHS COVID-19 service to manage post-COVID-19 syndrome in the community. Journal of Primary Care & Community Health, 12, 215013272110109. https://doi.org/10.1177/21501327211010994
2. Nalbandian, A., Sehgal, K., Gupta, A., Madhavan, M. V., McGroder, C., Stevens, J. S., et al. (2021). Post-acute COVID-19 syndrome. Nature Medicine, 27, 601–615. https://doi.org/10.1038/s41591-021-01283-z
3. Guido, C., Lucidi, F., Midulla, F., Zicari, A., Bove, E., Avenoso, F., et al. (2022). Neurological and psychological effects of long covid in a young population: A cross-sectional study. Frontiers in Neurology, 13. https://doi.org/10.3389/fneur.2022.925144
4. Davis, H. E., McCorkell, L., Vogel, J. M., Topol, E. J. Long COVID: major findings, mechanisms and recommendations. Nature Reviews Microbiology, 21(3), 133–46. https://doi.org/10.1038/s41579-022-00846-2
5. Kim, Y., Bae, S., Chang, H. H., Park, J., Cho, H., Lee, J. S., et al. (2023). Long COVID prevalence and impact on quality of life 2 years after acute COVID-19. Scientific Reports, 13, 11207. https://doi.org/10.1038/s41598-023-36995-4
6. Woodrow, M., Carey, C., Ziauddeen, N., Thomas, R., Akrami, A., Lutje, V., et al. (2023). Systematic review of the prevalence of Long COVID. Open Forum Infectious Diseases, 10(7), ofad233. https://doi.org/10.1093/ofid/ofad233
7. Forstag, E. H., Denning, L. A. Overview of Long COVID and Disability. (2022, June 30). Long COVID - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK585192/ Accessed: 10.09.2024.
8. Goërtz, Y. M. J., Van Herck, M., Delbressine, J. M., Vaes, A. W., Meys, R., Machado, F. V. C., et al. (2020). Persistent symptoms 3 months after a SARS-CoV-2 infection: the post-COVID-19 syndrome?. ERJ Open Research, 6(4), 00542-2020. https://doi.org/10.1183/23120541.00542-2020
9. Stefanou, M., Palaiodimou, L., Bakola, E., Smyrnis, N., Papadopoulou, M., Paraskevas, G. P., et al. (2022). Neurological manifestations of long-covid syndrome: a narrative review. Therapeutic Advances in Chronic Disease, 13, 204062232210768. https://doi.org/10.1177/20406223221076890
10. Iqbal, F., Lam, K., Sounderajah, V., Clarke, J., Ashrafian, H., Darzi, A. (2021). Characteristics and predictors of acute and chronic post-covid syndrome: a systematic review and meta-analysis. EClinicalMedicine, 36, 100899. https://doi.org/10.1016/j.eclinm.2021.100899
11. Maggi, G., Baldassarre, I., Barbaro, A., Sperduti, I., Sciarretta, A., Mongelli, F., et al. (2021). Age- and gender-related differences in the evolution of psychological and cognitive status after the lockdown for the COVID-19 outbreak: A follow-up study. Neurological Sciences, 43, 1521-1532. https://doi.org/10.1007/s10072-021-05768-0
12. Morin, C. M., Carrier, J., Bastien, C., & Godbout, R. (2020). Sleep and circadian rhythm in response to the COVID-19 pandemic. Canadian Journal of Public Health, 111(5), 654-657.

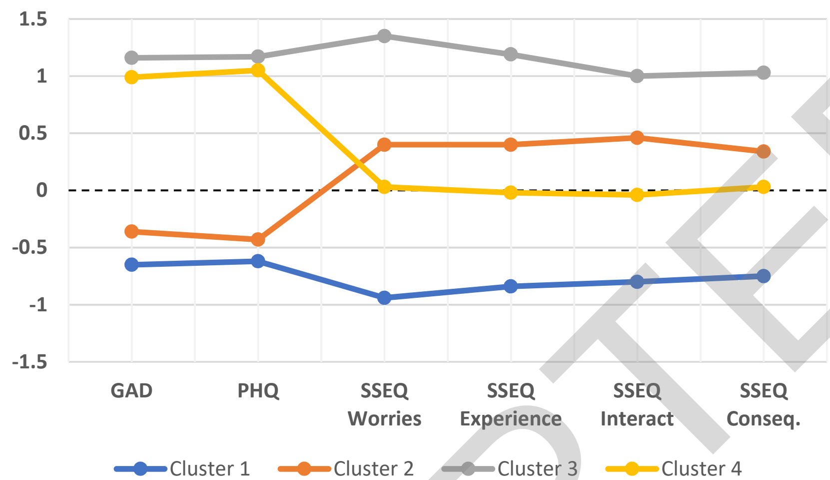
13. Stefano, G. B. (2021). Historical insight into infections and disorders associated with neurological and psychiatric sequelae similar to long covid. Medical Science Monitor, 27. https://doi.org/10.12659/msm.931447
14. Choutka, J., Jansari, V., Hornig, M. et al. Unexplained post-acute infection syndromes. Nat Med 28, 911–923 (2022). https://doi.org/10.1038/s41591-022-01810-6
15. van Kessel, S. A. M., Olde Hartman, T. C., Lucassen, P. L. B. J., & van Jaarsveld, C. H. M. (2022). Post-acute and long-COVID-19 symptoms in patients with mild diseases: a systematic review. Family practice, 39(1), 159–167. https://doi.org/10.1093/fampra/cmab076
16. Wong-Chew, R. M., Rodríguez Cabrera, E. X., Rodríguez Valdez, C. A., Lomelin-Gascon, J., Morales-Juárez, L., de la Cerda, M. L. R., et al. (2022). Symptom cluster analysis of long COVID-19 in patients discharged from the Temporary COVID-19 Hospital in Mexico City. Therapeutic Advances in Infectious Disease, 9, 20499361211069264. https://doi.org/10.1177/20499361211069264
17. Kisiel, M. A., Lee, S., Malmquist, S., Rykatkin, O., Holgert, S., Janols, H., et al. (2023). Clustering analysis identified three long COVID phenotypes and their association with general health status and working ability. Journal of Clinical Medicine, 12(11), 3617. https://doi.org/10.3390/jcm12113617
18. Zhang, H., Zang, C., Xu, Z., et al. (2023). Data-driven identification of post-acute SARS-CoV-2 infection subphenotypes. Nature Medicine, 29, 226–235. https://doi.org/10.1038/s41591-022-02116-3
19. Kenny, G., McCann, K., O'Brien, C., Savinelli, S., Tinago, W., Yousif, O., et al. (2022). Identification of distinct long COVID clinical phenotypes through cluster analysis of self-reported symptoms. Open Forum Infectious Diseases, 9(4), ofac060. https://doi.org/10.1093/ofid/ofac060
20. Fernández-de-Las-Peñas, C., Martín-Guerrero, J. D., Florencio, L. L., Navarro-Pardo, E., Rodríguez-Jiménez, J., Torres-Macho, J., et al. (2023). Clustering analysis reveals different profiles associating long-term post-COVID symptoms, COVID-19 symptoms at hospital admission, and previous medical co-morbidities in previously hospitalized COVID-19 survivors. Infection, 51(1), 61–69. https://doi.org/10.1007/s15010-022-01822-x
21. Fancourt, D., Steptoe, A., & Bu, F. (2023). Psychological consequences of long COVID: Comparing trajectories of depressive and anxiety symptoms before and after contracting SARS-CoV-2 between matched long- and short-COVID groups. The British Journal of Psychiatry, 222(2), 74–81. https://doi.org/10.1192/bjp.2022.155
22. Fernández-de-las-Peñas, C., Martín-Guerrero, J. D., Navarro-Pardo, E., Rodríguez-Jiménez, J., Torres-Macho, J., & Pellicer-Valero, Ó. J. (2022). Clustering analysis reveals different profiles associating long-term post-covid symptoms, covid-19 symptoms at hospital admission and previous medical co-morbidities in previously hospitalized covid-19 survivors. Infection, 51(1), 61-69. https://doi.org/10.1007/s15010-022-01822-x
23. Goldhaber, N. H., Kohn, J. N., Ogan, W. S., Sitapati, A., Longhurst, C. A., Wang, A., et al. (2022). Deep dive into the long haul: Analysis of symptom clusters and risk factors for post-acute sequelae of COVID-19 to inform clinical care. International Journal of Environmental Research and Public Health, 19(24), 16841. https://doi.org/10.3390/ijerph192416841

24. Tsuchida, T., Yoshimura, N., Ishizuka, K., Katayama, K., Inoue, Y., Hirose, M., et al. (2023). Five cluster classifications of long COVID and their background factors: A cross-sectional study in Japan. Clinical and Experimental Medicine. https://doi.org/10.1007/s10238-023-01057-6
25. Taquet, M., Geddes, J. R., Husain, M., Luciano, S., & Harrison, P. J. (2021). 6-month neurological and psychiatric outcomes in 236,379 survivors of COVID-19: A retrospective cohort study using electronic health records. The Lancet Psychiatry, 8(5), 416–427. https://doi.org/10.1016/S2215-0366(21)00084-5
26. Caspersen, I. H., Magnus, P., & Trogstad, L. (2022). Excess risk and clusters of symptoms after COVID-19 in a large Norwegian cohort. European Journal of Epidemiology. https://doi.org/10.1007/s10654-022-00847-8
27. Re'em, Y., Stelson, E. A., Davis, H. E., et al. (2023). Factors associated with psychiatric outcomes and coping in Long COVID. Nature Mental Health, 1, 361–372. https://doi.org/10.1038/s44220-023-00064-6
28. Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M. R., et al. (2007). No health without mental health. The Lancet, 370, 859–77. https://doi.org/10.1016/S0140-6736(07)61238-0
29. Seidl, E., & Seidl, O. (2021). Do my feelings fit the diagnosis? Avoiding misdiagnoses in psychosomatic consultation services. Journal of Healthcare Risk Management, 41, 9–17.
30. Alwan, N. A. (2022). Lessons from Long COVID: Working with patients to design better research. Nature Reviews Immunology, 22, 201–202.
31. Roth, P. H., & Gadebusch-Bondio, M. (2022). The contested meaning of 'long COVID'—patients, doctors, and the politics of subjective evidence. Social Science & Medicine, 292, 114619.
32. Pantelic, M., & Alwan, N. (2021). The stigma is real for people living with long COVID. British Medical Journal. https://blogs.bmj.com/bmj/2021/03/25/marija-pantelic-and-nisreen-alwan-the-stigma-is-real-for-people-living-with-long-covid/ Accessed: 10.09.2024.
33. WHO. Post COVID-19 condition (long COVID) (2022). Available at: https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition Accessed: 10.09.2024.
34. Gómez Bravo, R., Infanti, A., Billieux, J., Ritzen, M., Psy-Long-COVID Consortium, Vögele, C., et al. (2023). The psychological syndrome associated with Long COVID: A study protocol. Frontiers in Epidemiology, 3, 1193369. https://doi.org/10.3389/fepid.2023.1193369
35. Global Initiative on Sharing All Influenza Data. Available at: https://ourworldindata.org/grapher/covid-variants-bar?country=FRA~LUX~DEU~CHE Accessed: 10.09.2024.
36. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis. A global perspective. 7th ed. Upper Saddle River (N.J.): Pearson Education.
37. Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. Journal of Statistical Software, 61(6), 1–36. https://doi.org/10.18637/jss.v061.i06 Accessed: 18.07.2023.
38. Ebbert, D. (2019). chisq.posthoc.test: A post hoc analysis for Pearson's chi-squared test for count data. R package version 0.1.2. Available at: https://CRAN.Rproject.org/package=chisq.posthoc.test Accessed: 18.07.2023.

39. Sun, J., He, W. T., Wang, L., Lai, A., Ji, X., Zhai, X., et al. (2020). COVID-19: Epidemiology, evolution, and cross-disciplinary perspectives. Trends in Molecular Medicine, 26(5), 483–495.
40. Sivan, M., Smith, A. B., Osborne, T., Goodwin, M., Lawrence, R. R., Baley, S., et al. (2024). Long COVID Clinical Severity Types Based on Symptoms and Functional Disability: A Longitudinal Evaluation. Journal of Clinical Medicine, 13(7), 1908. https://doi.org/10.3390/jcm13071908

Figure 1. **Clustering solutions. A) Affinity Propagation. B) Kmeans.**

(a) AffinityPropagation

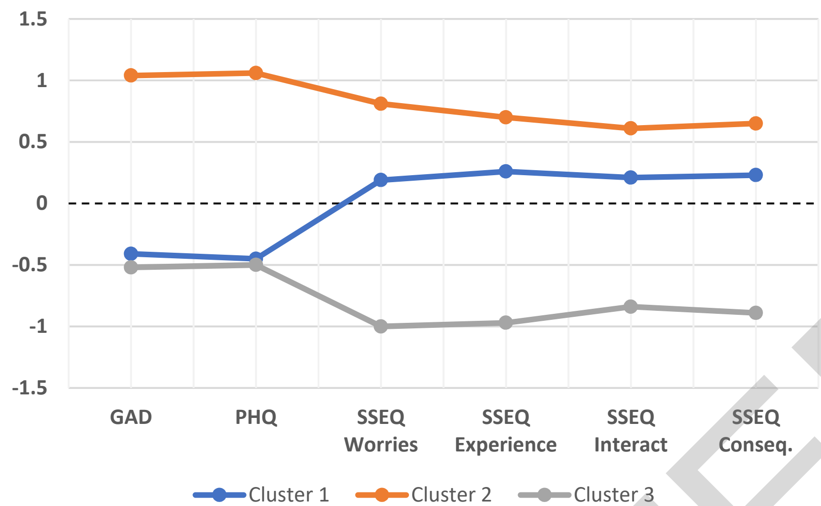


	GAD (SD)	PHQ (SD)	SSEQ Worries (SD)	SSEQ Experience (SD)	SSEQ Interact (SD)	SSEQ Conseq. (SD)
Cluster 1 (N = 423)	-0.65 (0.72)	-0.62 (0.64)	-0.94 (0.65)	-0.84 (0.78)	-0.8 (0.65)	-0.75 (0.89)
Cluster 2 (N = 399)	-0.36 (0.68)	-0.43 (0.58)	0.4 (0.6)	0.4 (0.70)	0.46 (0.78)	0.34 (0.68)
Cluster 3 (N = 226)	1.16* (0.61)	1.17* (0.79)	1.35 (0.62)	1.19 (0.45)	1 (0.82)	1.03 (0.62)
Cluster 4 (N = 170)	0.99* (0.57)	1.05* (0.65)	0.03 (0.64)	-0.02 (0.77)	-0.04 (0.84)	0.03 (0.86)

* Clusters are not significantly different (column)

Silhouette: 0.36
Calinski-Harabasz score: 424.15
Davies bouldin: 1.42

(b) Kmeans



	GAD (SD)	PHQ (SD)	SSEQ Worries (SD)	SSEQ Experience (SD)	SSEQ Interact (SD)	SSEQ Conseq. (SD)
Cluster 1 (N = 455)	-0.41* (0.66)	-0.45* (0.59)	0.19 (0.63)	0.26 (0.7)	0.21 (0.82)	0.23 (0.66)
Cluster 2 (N = 389)	1.04 (0.62)	1.06 (0.76)	0.81 (0.79)	0.7 (0.76)	0.61 (0.93)	0.65 (0.77)
Cluster 3 (N = 374)	-0.52* (0.86)	-0.5* (0.78)	-1 (0.64)	-0.97 (0.72)	-0.84 (0.64)	-0.89 (0.89)

* Clusters are not significantly different (column)

Silhouette: 0.35
Calinski-Harabasz score: 513.12
Davies bouldin: 1.46

Table 1: Long COVID symptoms (answer = yes and still have)

	Cluster 1 n (%)		Cluster 2 n (%)		Cluster 3 n (%)		Cluster 4 n (%)		Chi-square test			Cramer's V
	N	%	N	%	N	%	N	%	df	X^2	p	
Loss of taste/smell	106^{a}	25.1	99^{a}	24.8	41^{a}	24.1	70^{a}	31	6	7.996	.238	.057
Other ear, nose and throat disorder	101^{a}	23.9	143^{b}	35.8	72^{b}	42.4	$71^{a,b}$	21.4	6	29.184	< .001	.109
Headache, migraines	221^{a}	52.2	$284^{b,c}$	71.2	133^{c}	78.2	$141^{a,b}$	62.4	6	54.062	< .001	.149
Memory or concentration	328^{a}	77.5	344^{b}	86.2	157^{b}	92.4	198^{b}	87.6	6	33.398	< .001	.117
Other neurological symptoms	225^{a}	53.2	288^{b}	72.2	150^{c}	88.2	163^{b}	72.1	6	83.537	< .001	.185
Cardiorespiratory	207^{a}	48.9	259^{b}	64.9	124^{b}	72.9	149^{b}	65.9	6	44.754	< .001	.136
Gastrointestinal symptoms	70^{a}	16.5	114^{b}	28.6	56^{b}	32.9	$53^{a,b}$	23.5	6	40.407	< .001	.129
Vascular/lymph node symptoms	114^{a}	27	164^{b}	41.1	92^{c}	54.1	$82^{a,b}$	36.3	6	51.745	< .001	.146
Skin symptoms	58^{a}	13.7	102^{b}	25.6	41^{b}	24.1	$40^{a,b}$	17.7	6	20.901	.002	.093
Other	231^{a}	54.6	282^{b}	70.7	121^{b}	71.2	$137^{a,b}$	60.6	6	34.058	< .001	.118

Each subscript letter denotes a subset of clusters categories whose column proportions do not differ significantly from each other at the .05 level.

Table 2: Description of symptom clusters and associated characteristics.

Clusters	Characteristics
Cluster 1:	Characterized by fewer somatic symptoms and more likely to be in a partnership, which could correlate with lesser psychological symptoms.
Cluster 2:	Displays a varied pattern of mental distress, with a notable presence of somatoform symptomatology, leading to significant functional and professional challenges.
Cluster 3:	Exhibits pronounced physical issues, a higher incidence of pre-existing mental health conditions, and a substantial impact on daily life, indicating severe psychological symptoms as well.
Cluster 4:	Presents a varied pattern of mental distress too, but with a stronger presence of anxiety and depression, impacting interpersonal and social interactions.