

## ARTICLE



# The phenotype and prediction of long-term physical, mental and cognitive COVID-19 sequelae 20 months after recovery, a community-based cohort study in China

Yimiao Zhao<sup>1,2,8</sup>, Le Shi<sup>3,8</sup>, Zhendong Jiang<sup>4,8</sup>, Na Zeng<sup>1,2</sup>, Huan Mei<sup>1,2</sup>, Yu Lu<sup>4</sup>, Jinfeng Yang<sup>4</sup>, Fenshu Jin<sup>4</sup>, Shuyu Ni<sup>1,2</sup>, Shuilin Wu<sup>1,2</sup>, Sizhen Su<sup>3</sup>, Yongbo Zheng<sup>3</sup>, Kai Yuan<sup>3</sup>, Wei Yan<sup>3</sup>, Siwei Sun<sup>3</sup>, Shiqiu Meng<sup>1</sup>, Jie Sun<sup>5</sup>, Zunera Khan<sup>6</sup>, Dag Aarsland<sup>6</sup>, Jie Shi<sup>1</sup>, Xiangyou Li<sup>4</sup>, Yanping Bao<sup>1,2</sup> and Lin Lu<sup>1,3,7</sup>

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Long-term sequelae clustering phenotypes are important for precise health care management in COVID-19 survivors. We reported findings for 1000 survivors 20 months after diagnosis of COVID-19 in a community-based cohort in China. Sequelae symptoms were collected from a validated questionnaire covering 27 symptoms involved in five organ systems including self-reported physical condition, dyspnea, cognitive function and mental health. The generalized symptoms were reported with the highest rate (60.7%), followed by the mental (48.3%), cardiopulmonary (39.8%), neurological (37.1%; cognitive impairment, 15.6%), and digestive symptoms (19.1%). Four clusters were identified by latent class analysis: 44.9% no or mild group (cluster 1), 29.2% moderate group with mainly physical impairment (cluster 2), 9.6% moderate group with mainly cognitive and mental health impairment (cluster 3), and 16.3% severe group (cluster 4). Physical comorbidities or history of mental disorders, longer hospitalization periods and severe acute illness predicted severe group. For moderate group, adults less than 60 years, with physical comorbidities and severe acute illness were more likely to have physical symptoms, while adult women with longer hospitalization stays had increased risk of cognitive and mental health impairment. Overall, among more than half of community COVID-19 survivors who presented moderate or severe sequelae 20 months after recovery, three-tenth had physical vulnerability that may require physical therapy aiming to improve functioning, one-tenth mental or cognitive vulnerable cases need psychotherapy and cognitive rehabilitation, and one-sixth severe group needs multidisciplinary clinical management. The remaining half is free to clinical intervention. Our findings introduced an important framework to map numerous symptoms to precise classification of the clinical sequelae phenotype and provide information to guide future stratified recovery interventions.

*Molecular Psychiatry* (2023) 28:1793–1801; <https://doi.org/10.1038/s41380-023-01951-1>

## INTRODUCTION

As of June, 2022, coronavirus disease 2019 (COVID-19) has resulted in more than 500 million confirmed cases with 6 million deaths in more than 200 countries and regions [1]. With the increasing number of patients recovering from COVID-19, the long-term consequences of SARS-CoV-2 infection are now being highlighted. The evolving epidemiological and clinical evidence found that a proportion of survivors report ongoing health problems, which are described by a series of terms, such as long COVID-19, long-haul COVID-19 or post-acute COVID-19 syndrome [2].

A number of studies revealed that COVID-19 had long-term effects on multiple organ systems including pulmonary, cardiovascular, neurological, gastrointestinal, psychiatric and dermatologic systems [3–10]. According to a meta-analysis of sequelae for COVID-19, more than half of survivors have reported at least one

sequelae symptom for up to 12 months after the acute phase [7]. The most common findings included abnormalities on lung CT, abnormal pulmonary function tests, generalized symptoms, psychiatric symptoms, and neurological symptoms mainly cognitive deficits and memory impairment. In a two-year hospital-based follow-up study of COVID-19 patients, 55% reported at least one sequelae symptom, with fatigue or muscle weakness, sleep disturbances and smell disorder always being the common symptoms [11]. There is growing concern about brain damage caused by COVID-19. The recent definition of post-COVID-19 from WHO includes fatigue, dyspnea and cognitive dysfunction as three most common symptoms [12]. Moreover, stressful medical care, long hospital stays with numerous restrictions on mobility and social isolation, and stigma can contribute a few people to experiencing the lingering mental symptoms including

<sup>1</sup>National Institute on Drug Dependence and Beijing Key Laboratory of Drug Dependence, Peking University, Beijing 100191, China. <sup>2</sup>School of Public Health, Peking University, Beijing 100191, China. <sup>3</sup>Institute of Mental Health, National Clinical Research Center for Mental Disorders, Key Laboratory of Mental Health and Peking University Sixth Hospital, Peking University, Beijing 100191, China. <sup>4</sup>Wuhan Wuchang Hospital, Wuhan University of Science and Technology, Wuhan, China. <sup>5</sup>Pain Medicine Center, Peking University Third Hospital, Beijing 100191, China. <sup>6</sup>Department of Old Age Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. <sup>7</sup>Peking-Tsinghua Center for Life Sciences and PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100191, China. <sup>8</sup>These authors contributed equally: Yimiao Zhao, Le Shi, Zhendong Jiang. ✉email: [lixiangyou@wcykyz.cn](mailto:lixiangyou@wcykyz.cn); [baoy@bjmu.edu.cn](mailto:baoy@bjmu.edu.cn); [linlu@bjmu.edu.cn](mailto:linlu@bjmu.edu.cn)

Received: 25 September 2022 Revised: 24 December 2022 Accepted: 6 January 2023

Published online: 23 January 2023

depression, anxiety, insomnia and post-traumatic stress disorder (PTSD) symptoms [3–5, 7, 13–18].

In response to the COVID-19 survivors who experience a wide array of complex, multifactorial symptoms, precise classification for COVID-19 sequelae is integral part of the rehabilitation therapy, which will guide health care practitioners support targeted interventions for survivors based on sequelae phenotype. Few studies identified clinical phenotype of long COVID with different clustering methods. A UK multicenter cohort study of survivorship of hospitalized patients with COVID-19 just identified four recovery clusters by cluster analysis, but no further investigated associated factors with recovery clusters [4]. A UK community study identified a subset of COVID-19 survivors with predominantly respiratory symptoms and another subset of patients with tiredness [19]. Another study from Spain conducted clustering analysis among not only sequelae symptoms but also some clinical characteristics like medical co-morbidities and prevalence of symptoms at hospital admission [20]. Several studies only focused on symptomology without characterizing sequelae phenotype. The longest cohort of COVID-19 sequelae lasted up to two years after discharge, but it was based on a hospital which was specifically built for COVID-19 treatment [11]. The participants had some special characteristics for this hospital. There is an urgent need for community-based follow up study to recruit COVID-19 survivors from a wider range of diverse type of hospital sources, aiming to explore the general feature and phenotype of long-term consequences of COVID-19 for precise classification and intervention. We conducted a community-based cohort study, to identify sequelae phenotype and predict factors associated with clusters, which will become a foundation of evidenced based information guiding future tailored recovery interventions.

## METHODS

### Study design and participants

We recruited COVID-19 survivors from the communities in Hongshan district, Wuhan City, Hubei Province from October 12 to November 19, 2021. We invited all participants with laboratory confirmed or clinician-diagnosed COVID-19 from their medical records and the patients lists of health checkup was obtained from corresponding management agency of Hongshan District by the Wuchang Hospital. We excluded participants under 15 years old in this analysis. We provided participants with an electronic or paper questionnaire covering physical, cognitive and mental health. Most participants received the Montreal Cognitive Assessment (MoCA) to assess cognitive function. The results of the antibody and nucleic acid tests were collected from their medical records.

This study received approval from the ethics committee of Peking University Sixth Hospital (Institute of Mental Health). This study followed the American Association for Public Opinion Research (AAPOR) reporting guidelines and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement (appendix pp 2–4). Informed consent was obtained from all study participants prior to all study procedures.

### Measurement of long-term COVID-19 sequelae

Demographic characteristics included sex, age, occupation and education. Smoking, drinking, COVID-19 vaccination status, physical comorbidities, a history of mental disorders and a family history of mental disorders were also collected. According to the Chinese Clinical Guidance for COVID-19 Pneumonia Diagnosis and Treatment, the disease severity was classified three types: mild (i.e., the clinical symptoms are mild and there was no sign of pneumonia on chest imaging), moderate (the clinical symptoms are fever and respiratory symptoms, and radiologic assessments found signs of pneumonia) and severe/critical (severe, patients meet any of the following conditions: (1) shortness of breath,  $RR \geq 30$  times/min; (2) oxygen saturation  $\leq 93\%$  at rest; (3) alveolar oxygen partial pressure/fraction of inspiration  $O_2$  ( $PaO_2/FiO_2$ )  $\leq 300$  mmHg; (4) pulmonary imaging with significant progression of lesion  $> 50\%$  within 24–48 h; critical, patients meet any of the following conditions: (1) respiratory failure requiring mechanical ventilation; (2) shock; (3) combined with other organ failure needed ICU monitoring and treatment) [21].

27 sequelae symptoms from five organ systems (i.e., generalized, neurological, mental, cardiopulmonary and digestive systems) were collected from participants. Among them, 21 self-reported symptoms were evaluated by the symptom questionnaire, and the remaining 6 symptoms were measured by validated scales. The contents of the symptom questionnaire and assessment tools of outcomes were shown in appendix (Supplementary Table S1). Specifically, dyspnea was assessed by the modified British Medical Research Council (mMRC) dyspnea scale, which is a five-category scale to characterize the level of dyspnea with physical activity in which higher scores correspond with increased dyspnea, and the total scores higher than 0 were considered as dyspnea [22]. We also assessed quality of life used by the EuroQol five-dimension three level (EQ-5D-3L) questionnaire and the EuroQol Visual Analogue Scale (EQ-VAS) [23].

Cognitive function was measured by objective assessment via the validated Chinese version of Montreal Cognitive Assessment (MoCA) [24] and subjective assessment via the abbreviated Cognitive Failure Questionnaire-14 (CFQ-14) [25]. The MoCA evaluates global cognitive function with a total score of 0–30, which was interpreted as follows: normal cognitive function (26–30), mild (18–25), moderate (10–17) and severe (0–9) cognitive impairment. The CFQ-14 measures daily life cognitive failures, which has 14 questions with a 5-point Likert scale resulting in a factored score ranging from 0–100, with a score of  $\geq 43$  indicating cognitive failures [26]. The survivors with the MoCA scores lower than 18 (moderate or severe cognitive impairment) or the CFQ-14 scores higher than 43 (cognitive failure) were defined as cognitive impairment.

Mental health included depression symptoms measured by the Patient Health Questionnaire-9 (PHQ-9) [27], anxiety symptoms measured by the Generalized Anxiety Disorder-7 (GAD-7) [28], insomnia symptoms measured by the Insomnia Severity Index (ISI) [29], and PTSD symptoms measured by the PTSD checklist for DSM-5 (PCL-5) [30]. The total scores of these scales were interpreted as follows: PHQ-9, normal (0–4), mild (5–9), moderate to severe (10–27) depression symptoms; GAD-7, normal (0–4), mild (5–9), moderate to severe (10–21) anxiety symptoms; ISI, normal (0–7), subthreshold (8–14), moderate to severe (15–28) insomnia symptoms. We used the following cut-off scores as having depression (PHQ-9  $> 4$ ), anxiety (GAD-7  $> 4$ ), insomnia (ISI  $> 7$ ) or PTSD symptoms (PCL-5  $\geq 33$ ).

### Statistical analysis

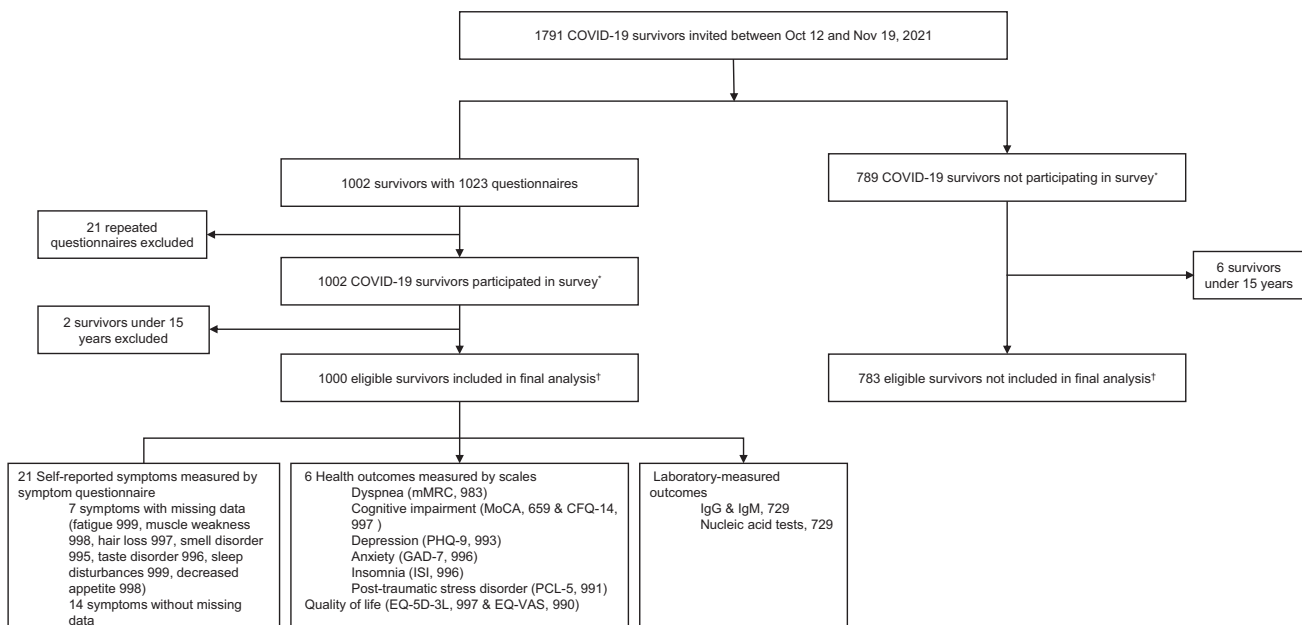
Continuous variables were presented as mean (SD) or median (IQR). Binary and categorical variables were presented as counts and percentages. Characteristics of participants were stratified by the severity of the acute illness. A chi-square test was used to identify differences in proportions across multiple categories. For normally distributed continuous data, one-way analysis of variance (ANOVA) was used to test differences across categories, with Kruskal–Wallis tests used for non-normally distributed data.

To identify the sequelae clusters of the COVID-19 survivors, unsupervised clustering of self-reported symptoms, and dyspnea, cognitive function and mental health evaluated by scales was undertaken using latent class analysis. The final number of classes were determined based on the conceptual meaning, smallest estimated class proportions, and statistical model fit indices, such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), adjusted BIC (aBIC) and entropy [31]. The Lo-Mendell-Rubin (LMR) test is also used to determine the number of classes in latent class analysis [32]. This is obtained by the k class run also doing a k-1 class analysis and using the derivatives from models to compute the p-value. A low p-value rejects the k-1 class model in favor of the k class model. To explore the predictive factors of sequelae phenotype, we used multinomial logistic regression for the sequelae clusters derived from latent class analysis. We adjusted for sex, age, physical comorbidities, history of mental disorders, duration of hospitalization and acute disease severity. All tests were two-tailed and p values of less than 0.05 were considered statistically significant. Latent class analysis was done with Mplus version 8.3. Other analyses used Stata MP version 16.

## RESULTS

### Participants

We invited 1791 COVID-19 survivors to take part in our survey from the communities, with 1783 survivors aged older than 15 years eligible for our survey. 1023 questionnaires were received and 1000 survivors were included in final analysis after excluding 21 repeated questionnaires and 2 participants under 15 years,



**Fig. 1 Flow chart of the study.** \*1791 COVID-19 survivors invited in our survey, with 1002 survivors willing to participate in survey and 789 not participating in survey. †1783 COVID-19 survivors eligible for survey with the exclusion of 8 survivors under 15 years (1000 eligible survivors included in analysis and 783 eligible survivors not included). The response rate: 1000 eligible survivors included in analysis / 1783 eligible survivors invited = 56.1%. The effective rate: 1000 eligible questionnaires/1023 completed questionnaires = 97.8%.

with the response rate of 56.1% and the effective questionnaire rate of 97.8% (Fig. 1). The characteristics of COVID-19 survivors included in the final analysis were similar to those not included (Supplementary Table S4).

The mean age of the participants was 55.9 (13.8) years, with 446 (44.6%) men. The majority of survivors had a high school education or lower (high school, 32.3%; middle school, 21.7%; primary school or lower, 7.3%). 514 (51.4%) had at least one physical comorbidity and the most common physical comorbidity were hypertension (27.5%), fatty liver (15.7%), hyperlipemia (12.1%) and diabetes (10.4%). 43 patients (4.3%) had history of mental disorders and 10 (1.0%) had family history of mental disorders. 85 patients (8.5%) were diagnosed as severe or critical cases in the acute phase. The median length of hospital stay was 31.0 days (21.0–48.0). The participants were assessed at a median of 625.0 days (615.0–634.0) after COVID-19 diagnosis. Seropositivity of IgG was observed in most patients (94.9%), but a few of people had the seropositivity of IgM (2.2%). No reinfection of SARS-COV-2 was observed according to the nucleic acid test. 77.3% of participants have received at least one dose of a COVID-19 vaccine (Table 1).

### Phenotype of four sequelae clusters

Among 1000 participants, 27 symptoms from five system were classified by latent class analysis (Supplementary Table S2), we identified four sequelae clusters: 44.9% individuals were classified as no or mild group (cluster 1), which included the survivors reporting 1 (median) symptom with the lowest probabilities of physical and mental symptoms. 29.2% were classified as moderate symptoms with mainly physical impairment with 4.0 (median) symptoms (cluster 2) and 9.6% were classified as moderate symptoms with mainly cognitive and mental health impairment with 5.0 (median) symptoms (cluster 3). Both cluster 2 and cluster 3 had a moderate proportion of sequelae symptoms, but physical symptoms (generalized, 95.2%; cardiopulmonary, 49.7%; digestive, 20.9%) were more common in cluster 2, and neurological (32.3%) especially cognitive impairment (25.0%) and mental symptoms (100.0%) were more common in cluster 3. 16.3% were classified as severe group (cluster 4), which had 11.0 (median) symptom with a greater proportion of physical (generalized, 100.0%; cardiopulmonary, 89.0%; digestive, 73.6%),

neurological (any, 92.6%; cognitive impairment, 40.5%) and mental symptoms (98.2%) (Fig. 2 & Fig. 3).

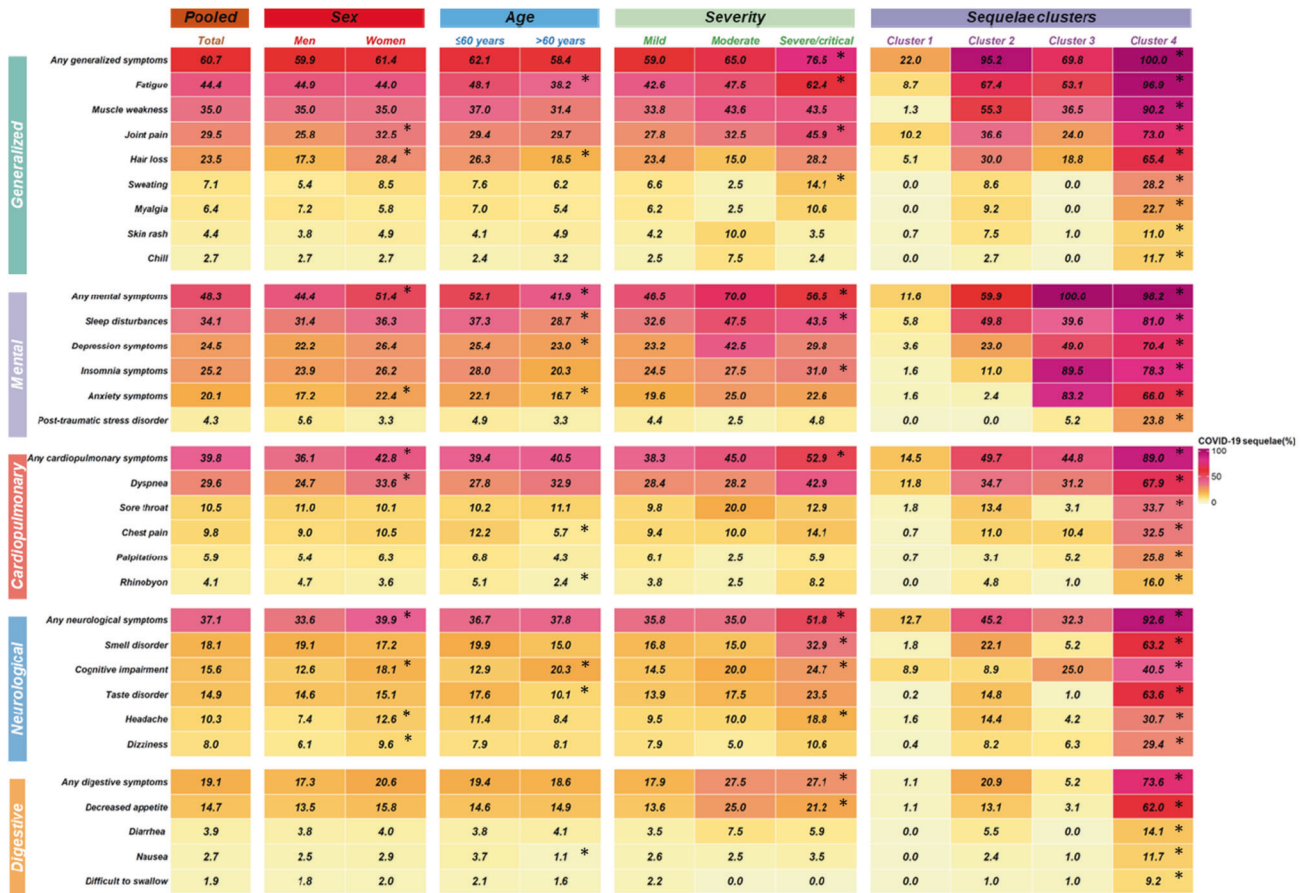
The demographic and clinical characteristics were different between the four sequelae clusters (Supplementary Table S3). The severe cluster 4 had a greater proportion of having at least one physical comorbidity (65.6%) and history of mental disorders (11.7%) than the other three clusters. All of individuals without any persistent symptoms were in the no or mild cluster. The median of symptom count (11.0 [10.0–14.0]) in the severe cluster was significantly higher than mild (1.0 [0.0–1.0]) and moderate group (mainly for physical impairment, 4.0 [3.0–6.0]; mainly for cognitive and mental health impairment, 5.0 [4.0–6.5]). The total duration of hospitalization was longer in the severe cluster (Cluster 4, 38.0 [23.0–57.0]) than the other clusters. There were higher rates of reporting some or extreme problems for mobility (23.5%), self-care (7.4%), usual activities (11.7%), pain/discomfort (74.1%) and anxiety/depression (60.5%) evaluated by the EQ-5D-3L and the scores of the EQ-VAS in the moderate cluster with mainly for cognitive and mental health impairment (66.5 [56.5–80.0]) and severe cluster (67.0 [50.5–75.0]) were lower than the other two clusters.

### Predictive factors for sequelae clusters

We explored the associated risk factors for the moderate and severe cluster. Compared with the mild cluster, participants in the moderate cluster with mainly physical impairment were more likely to be younger than 60 years (OR 1.43, 95% CI 1.03–1.98), to have at least one physical comorbidity (OR 1.68, 95% CI 1.22–2.31), and to be severe or critical cases in the acute phase (OR 1.81, 95% CI 1.01–3.23). Individuals who were female (OR 1.69, 95% CI 1.05–2.73), younger than 60 years (OR 2.08, 95% CI 1.23–3.51), and had longer hospital stays ( $\geq 60$  days, OR 2.57 [1.34–4.93]) had a high proportion of being identified into the moderate cluster with mainly cognitive and mental impairment in comparison with the mild cluster. Individuals aged younger than 60 years (OR 2.10, 95% CI 1.38–3.19), having at least one physical comorbidity (OR 2.71, 95% CI 1.81–4.08), having a history of mental disorders (OR 6.79, 95% CI 2.70–17.08), having longer hospital stays ( $\geq 30$  days, OR 1.67 [1.09–2.56]);  $\geq 60$  days, OR 2.65 [1.53–4.59]) and being

**Table 1.** Characteristics of COVID-19 survivors included in analysis by severity of acute illness.

	<b>Mild (N = 875)</b>	<b>Moderate (N = 40)</b>	<b>Severe/critical (N = 85)</b>	<b>Total (N = 1000)</b>	<b>p-value</b>
Age, years	55.3 (14.1)	58.9 (12.3)	60.6 (10.9)	55.9 (13.8)	0.001
Sex					0.006
Men	392 (44.8%)	9 (22.5%)	45 (52.9%)	446 (44.6%)	
Women	483 (55.2%)	31 (77.5%)	40 (47.1%)	554 (55.4%)	
Occupation					0.001
Students	8 (0.9%)	0 (0.0%)	0 (0.0%)	8 (0.8%)	
Farmers	22 (2.5%)	1 (2.5%)	5 (6.0%)	28 (2.8%)	
Workers	70 (8.0%)	1 (2.5%)	10 (12.0%)	81 (8.1%)	
Managers	47 (5.4%)	4 (10.0%)	7 (8.4%)	58 (5.8%)	
Healthcare workers	21 (2.4%)	0 (0.0%)	0 (0.0%)	21 (2.1%)	
Retirees	381 (43.6%)	27 (67.5%)	51 (61.4%)	459 (46.1%)	
Business service staff	63 (7.2%)	1 (2.5%)	2 (2.4%)	66 (6.6%)	
Logistics staff	7 (0.8%)	0 (0.0%)	1 (1.2%)	8 (0.8%)	
Others	254 (29.1%)	6 (15.0%)	7 (8.4%)	267 (26.8%)	
Education					0.106
Primary school or lower	65 (7.5%)	4 (10.0%)	3 (3.6%)	72 (7.3%)	
Middle school	179 (20.6%)	12 (30.0%)	24 (28.9%)	215 (21.7%)	
High school	290 (33.4%)	6 (15.0%)	24 (28.9%)	320 (32.3%)	
College or higher	335 (38.6%)	18 (45.0%)	32 (38.6%)	385 (38.8%)	
Smoking					<0.001
Never-smoker	761 (87.1%)	35 (87.5%)	63 (75.0%)	859 (86.1%)	
Former smoker	73 (8.4%)	2 (5.0%)	19 (22.6%)	94 (9.4%)	
Current smoker	40 (4.6%)	3 (7.5%)	2 (2.4%)	45 (4.5%)	
Drinking					0.001
Never-drinker	765 (87.4%)	35 (87.5%)	63 (75.0%)	863 (86.4%)	
Former drinker	46 (5.3%)	1 (2.5%)	14 (16.7%)	61 (6.1%)	
Current drinker	64 (7.3%)	4 (10.0%)	7 (8.3%)	75 (7.5%)	
Physical comorbidities					
Any of the following physical diseases	432 (49.4%)	24 (60.0%)	58 (68.2%)	514 (51.4%)	0.002
Hypertension	224 (25.6%)	14 (35.0%)	37 (43.5%)	275 (27.5%)	0.001
Fatty liver	129 (14.7%)	8 (20.0%)	20 (23.5%)	157 (15.7%)	0.078
Hyperlipemia	101 (11.5%)	4 (10.0%)	16 (18.8%)	121 (12.1%)	0.133
Diabetes	84 (9.6%)	4 (10.0%)	16 (18.8%)	104 (10.4%)	0.029
Coronary heart diseases	64 (7.3%)	2 (5.0%)	9 (10.6%)	75 (7.5%)	0.456
Malignancy	15 (1.7%)	1 (2.5%)	1 (1.2%)	17 (1.7%)	0.863
Chronic liver disease	8 (0.9%)	2 (5.0%)	4 (4.7%)	14 (1.4%)	0.003
Chronic kidney disease	7 (0.8%)	0 (0.0%)	3 (3.5%)	10 (1.0%)	0.044
Chronic obstructive pulmonary disease	6 (0.7%)	1 (2.5%)	2 (2.4%)	9 (0.9%)	0.164
Immunodeficiency diseases	6 (0.7%)	0 (0.0%)	2 (2.4%)	8 (0.8%)	0.218
History of mental disorders	37 (4.2%)	0 (0.0%)	6 (7.1%)	43 (4.3%)	0.184
Family history of mental disorders	10 (1.1%)	0 (0.0%)	0 (0.0%)	10 (1.0%)	0.484
Total duration of hospitalization, days	29.0 (20.0, 47.0)	36.0 (23.0, 46.0)	44.0 (29.0, 64.5)	31.0 (21.0, 48.0)	<0.001
Time from diagnosis to follow-up, days	626.0 (615.0, 635.0)	617.0 (610.5, 632.5)	625.0 (617.0, 632.0)	625.0 (615.0, 634.0)	0.081
Vaccinated for COVID-19					0.022
Two doses	467 (53.4%)	28 (70.0%)	43 (50.6%)	538 (53.8%)	
One dose	199 (22.7%)	7 (17.5%)	29 (34.1%)	235 (23.5%)	
Not vaccinated	209 (23.9%)	5 (12.5%)	13 (15.3%)	227 (22.7%)	
Seropositivity					
IgG	572 (94.4%)	36 (94.7%)	84 (98.8%)	692 (94.9%)	0.218
IgM	16 (2.6%)	0 (0.0%)	0 (0.0%)	16 (2.2%)	0.190



**Fig. 2** Sequelae symptoms stratified by sex, age, severity of acute illness and sequelae clusters. \*means the p-value of subgroup test less than 0.05.

diagnosed as severe or critical cases in the acute phase (OR 2.29, 95% CI 1.20–4.39) were significantly higher risk to be in the severe cluster than that in the mild cluster (Table 2).

**DISCUSSION**

This study is one of the first to report the phenotype clusters of long-term sequelae of COVID-19 including the five organ systems among 1000 community-based survivors after 20 months recovery from COVID-19. We identified four clusters: nearly half of participants having no or mild impairment, nearly three-tenth of survivors having moderate impairment with mainly physical symptoms, nearly one-tenth suffering from moderate impairment with mainly cognitive and mental symptoms, and more than one-sixth suffering from severe impairment. Participants who had physical comorbidities or history of mental disorders, had longer hospital stays and were severe clinical illness cases in the acute phase with high risk to severe or moderate cluster. Physical comorbidities and long-term hospitalization stay were associated with the physical and mental or cognitive sequelae clusters. The findings indicate that the survivors in the four different clusters need different precise intervention and management strategy. Our findings introduced an important framework to map numerous symptoms to precise classification of the clinical sequelae phenotype and provided information for the medical resource reserve and allocation during the post pandemic era.

The different combination of characteristics of acute illness will predict the 4-phenotype classification. We identified that nearly half of survivors have no or mild sequelae symptoms, which aligned with the prevalence of mild cluster of the recovery clusters

derived from the Post-hospitalisation COVID-19 study (PHOSP-COVID) in the UK [4]. Our results revealed a certain percentage of patients suffering from moderate impairment with mainly cognitive function and mental health. This showed a possible relationship between SARS-CoV-2 infection and the development or deterioration of cognitive impairment and complaints. Regarding a wide range of cognitive assessment tools, different definitions of cognitive impairment and various population, the prevalence of cognitive impairment in COVID-19 survivors was varied. From an international online survey of individuals with suspected or confirmed COVID-19, 58.4% of patients reported cognitive dysfunction after six months [3]. Cognitive impairment was also found after other coronavirus infections, like SARS, and MERS [33], which can extrapolate the number of COVID-19 patients with neurological deficits. COVID-19 had a profound negative effect on the survivors’ mental health. Our study reported a high rate of any mental symptoms, which were the second common sequelae, and all of moderate clusters with mainly cognitive and mental health impairment had psychological symptoms. It is worth noting that even mildly affected patients can exhibit psychological symptoms, including depression, anxiety and insomnia symptoms. Another meta-analysis study for COVID-19 sequelae also demonstrated an overall mental symptoms prevalence of 19.7% [7]. We observed that both mental and neurological persisted up to 20 months, but a 2-year retrospective cohort study found the disparities in the trajectories of these two health conditions. The researchers demonstrated that compared to other respiratory diseases controls, the increased risk of mood and anxiety disorders in the COVID-19 convalescent phase were transient, but the risk of psychotic disorder, cognitive deficit, dementia, and epilepsy or



**Fig. 3 The rate of sequelae symptoms by four sequelae clusters.** **A** The rate of sequelae symptoms by four sequelae clusters. **B** The rate of five organ systems by four sequelae clusters. Cluster 1: no or mild group (no or mild physical, cognitive and mental health impairment) (44.9%). Cluster 2: moderate group with mainly physical impairment (29.2%). Cluster 3: moderate group with mainly cognitive and mental health impairment (9.6%). Cluster 4: severe group (severe physical, cognitive and mental health impairment) (16.3%).

seizures was still higher than controls at 2-year follow-up [34]. It is difficult for us to estimate the relative risk of long COVID due to the lack of controls, which may result to this difference. Several mechanisms have been presented to explain SARS-CoV-2's effects on the brain, including viral neurotropism, widespread systemic inflammation, and psychological burden of the pandemic across the world [35].

A few studies have reported that women infected with SARS-CoV-2 were more likely to get long lasting symptoms than men [5, 11], which is consistent with our findings that women had a higher percentage of sequelae and women were more likely to be classified into moderate clusters with mainly cognitive and mental health impairment. Sex differences in long COVID-19 syndrome may be explained by differences in immune system function between females and males. More rapid and robust immune responses can protect women from SARS-CoV-2 infection and

serious illness, but this condition may cause the long-lasting autoimmune diseases in the convalescent phase, which is related to the long COVID-19 [36, 37]. There is also a distinctive age distribution in COVID-19 survivors. According to our study, the risk of developing the lingering disorder is relatively lower in survivors aged less than 60 years than older people. There is no consensus on age differences in COVID-19 sequelae. Some studies reported that increased age was associated with more reports of long COVID-19 symptoms [5, 7], while some reported the opposite [4, 38, 39]. The PHOSP-COVID also found that middle age (40–59 years) was associated with not recovery from COVID, and their authors claimed in the 'news feature' section of Nature that these findings may be due to 'survivors bias' [4, 40]. We also speculated that older people were not willing to participate in the survey due to limited mobility and young people with more symptoms of long COVID were willing to accept assessment, which caused the selective bias and higher

**Table 2.** Factors that influence sequelae clusters according to multinomial logistic regression.

	Cluster 2 vs. Cluster 1, OR (95%CI)	p-value	Cluster 3 vs. Cluster 1, OR (95%CI)	p-value	Cluster 4 vs. Cluster 1, OR (95%CI)	p-value
Sex, women vs. men	1.06 (0.78, 1.45)	0.707	<b>1.69 (1.05, 2.73)</b>	<b>0.032</b>	1.48 (1.00, 2.18)	0.053
Age, ≤60 y vs. > 60 y	<b>1.43 (1.03, 1.98)</b>	<b>0.032</b>	<b>2.08 (1.23, 3.51)</b>	<b>0.006</b>	<b>2.10 (1.38, 3.19)</b>	<b>0.001</b>
Physical comorbidities, yes vs. no	<b>1.68 (1.22, 2.31)</b>	<b>0.001</b>	1.59 (0.99, 2.56)	0.055	<b>2.71 (1.81, 4.08)</b>	<b>&lt;0.001</b>
History of mental disorders, yes vs. no	1.82 (0.66, 4.99)	0.247	2.62 (0.73, 9.31)	0.138	<b>6.79 (2.70, 17.08)</b>	<b>&lt;0.001</b>
Total duration of hospitalization, days						
<30	Reference		Reference		Reference	
≥30	1.37 (0.98, 1.92)	0.065	1.58 (0.95, 2.62)	0.078	1.67 (1.09, 2.56)	0.019
≥60	1.53 (0.95, 2.45)	0.079	<b>2.57 (1.34, 4.93)</b>	<b>0.005</b>	<b>2.65 (1.53, 4.59)</b>	<b>&lt;0.001</b>
Disease severity						
Mild	Reference		Reference		Reference	
Moderate	1.58 (0.71, 3.50)	0.263	1.65 (0.56, 4.85)	0.360	1.57 (0.60, 4.14)	0.362
Severe/critical	<b>1.81 (1.01, 3.23)</b>	<b>0.045</b>	0.70 (0.23, 2.13)	0.535	<b>2.29 (1.20, 4.39)</b>	<b>0.012</b>

Bold values represent statistical significance  $p < 0.05$ .

proportion of long COVID observed in young people. However, according to a prospective cohort of home-isolated COVID-19 patients, young people with mild acute illness who didn't need a hospital treatment were also at risk of persistent symptoms of dyspnea and cognitive symptoms. Therefore, we still need to pay attention to long COVID in young people [41]. A positive association between comorbidities with increased risk of moderate impairment with mainly physical symptoms and severe impairment was observed in our study. Consistent with previous reports, comorbidities, such as diabetes, chronic cardiovascular or kidney disease, and cancer, are associated with both increased risk of severe acute illness and subsequent poor recovery [2, 4]. We found that the sequelae clusters were related to severity of acute illness, except for the moderate group with mainly cognitive and mental health impairment, further supporting the view that it is hard to attribute solely the long COVID-19 syndrome to the severity of the acute lung and other organs injury. Longer hospital stays, which not only represent disease severity, but also bring numerous restrictions on mobility and life style, may have affected mental health sequelae [7]. Moreover, the medical interventions for a long-term study in the hospital, especially mechanical ventilation, provided to alleviate conditions in acute infection, were also related to severe long-term sequelae symptoms.

We found that nearly three-tenth of survivors had moderate impairment with mainly physical symptoms, of which fatigue was the most common symptom. Similar results were observed in two-year follow-up long COVID-19 study, in which fatigue or muscle weakness was always the most frequent [11]. It is not surprising that some people infected with COVID-19 developed a debilitating chronic fatigue [42], because post-infectious fatigue syndromes are often preceded by several infectious agents, like SARS [43] and Ebola virus [44]. COVID-19 and other infectious diseases with lingering fatigue in the recovery may share a common pathogenic mechanism irrespective of the acute symptoms. In the moderate group with mainly physical symptoms, cardiopulmonary sequelae were observed in nearly half of survivors, and dyspnea was the most prevalent sequelae, indicating that it is necessary to pay attention to the long-term lung impairment and carry out serial objective assessments like lung CT scans, pulmonary function tests and six minute walk test (6MWT) during convalescence.

Our findings indicated that precise management and intervention strategies should be administrated for different sequelae phenotypes in community COVID-19 survivors. From identified risk factors of sequelae clusters, we can design specific management

for survivors in the convalescent phase according to their different demographic characteristics including age and sex, and clinical characteristics including medical comorbidities, duration of hospitalization and severity of acute illness. For survivors in the severe sequelae cluster, a multidisciplinary model of care should be supported on the basis of comprehensive evaluations, including physical therapy, occupational therapy, physical medicine and psychotherapy. For moderate group with mainly physical symptoms, physical therapy can be provided in the post-acute period to help survivors adapt and improve functioning, and return to normal life. For moderate group with mainly cognitive and mental health impairment, psychotherapy or behavioral interventions can cope with mental health difficulties, including depression, anxiety, insomnia and PTSD. Cognitive rehabilitation including retraining orientation, memory, attentions and executive functioning skills is also vital for this group of survivors with cognitive impairment [45–47].

The strengths of our community-based cohort study include assessing the comprehensive assessment of physical, cognitive and mental health outcomes in a large community-based sample of COVID-19 patients after acute infection, identifying sequelae clusters and investigating predictive factors for sequelae classification. However, this study has several limitations. Firstly, we cannot draw a clear conclusion that whether the sequelae symptoms were caused by SARA-CoV-2 infection or the premorbid diseases due to the lack of controls without COVID-19 diagnoses. Moreover, due to the absence of controls without COVID-19, we are difficult to compare prevalence of symptoms and estimate the relative risk of sequelae. Meanwhile, we need to validate the sequelae results in a large cohort study with healthy control in the future. Secondly, since most are self-reported symptoms, more objective assessments, like CT, MRI and PET are required to better characterize the long-term symptoms in future studies. Thirdly, despite the relatively low response rate and not random sampling, there is no difference between COVID-19 eligible survivors included and not included in final analysis and our participants are representative (Supplementary Table S4). Fourthly, it is unclear that how the sequelae clusters will progress during the recovery process, which needs to be investigated in further follow-up studies. Finally, information about treatment and condition during hospital stay, such as occurrence of delirium, ICU admission, use of mechanical ventilation and medical treatment were not available, which would have contributed to better interpretation of long-term health outcomes.

We identified four sequelae clusters after COVID-19 sequelae according to physical, cognitive and mental health symptoms, and investigated predictive factors associated with clusters. About half survivors need stratified and different degree and specific management and intervention 20 months after recovery from COVID-19. Our findings support the need to characterize clinical phenotypes according to validated assessment of overall health during convalescence, and suggest an important framework to map numerous symptoms to different clusters guiding future precise recovery interventions.

## REFERENCES

- World Health Organization. WHO Coronavirus (COVID-19) Dashboard. <https://covid19.who.int/>. Accessed 30 June 2022.
- Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. *Nat Med.* 2021;27:601–15.
- Davis HE, Assaf GS, McCorkell L, Wei H, Low RJ, Re'em Y, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine.* 2021;38:101019.
- Evans RA, McAuley H, Harrison EM, Shikotra A, Singapuri A, Sereno M, et al. Physical, cognitive, and mental health impacts of COVID-19 after hospitalisation (PHOSP-COVID): a UK multicentre, prospective cohort study. *Lancet Respir Med.* 2021;9:1275–87.
- Huang C, Huang L, Wang Y, Li X, Ren L, Gu X, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet.* 2021;397:220–32.
- Morin L, Savale L, Pham T, Colle R, Figueiredo S, Harrois A, et al. Four-month clinical status of a cohort of patients after hospitalization for COVID-19. *JAMA.* 2021;325:1525–34.
- Zeng N, Zhao YM, Yan W, Li C, Lu QD, Liu L et al. A systematic review and meta-analysis of long term physical and mental sequelae of COVID-19 pandemic: call for research priority and action. *Mol Psychiatry.* 2023;28:423–33.
- Brito J, Silva B, da Silva P, Cortez-Dias N, Silva D, Agostinho J, et al. Cardiovascular complications of COVID-19. *Heart Mind.* 2020;4:67–74.
- Wang J-A. Psychosomatic diseases and COVID-19. *Heart Mind.* 2021;5:101–2.
- Cereda A, Toselli M, Laricchia A, Mangieri A, Ruggiero R, Gallo F, et al. Stress-induced cardiomyopathy related to SARS-CoV-2. *Heart Mind.* 2020;4:57–58.
- Huang L, Li X, Gu X, Zhang H, Ren L, Guo L, et al. Health outcomes in people 2 years after surviving hospitalisation with COVID-19: a longitudinal cohort study. *Lancet Respir Med.* 2022;10:863–76.
- Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV. A clinical case definition of post-COVID-19 condition by a Delphi consensus. *Lancet Infect Dis.* 2022;22:e102–e107.
- Taquet M, Luciano S, Geddes JR, Harrison PJ. Bidirectional associations between COVID-19 and psychiatric disorder: retrospective cohort studies of 62 354 COVID-19 cases in the USA. *Lancet Psychiatry.* 2021;8:130–40.
- Liu L, Ni SY, Yan W, Lu QD, Zhao YM, Xu YY, et al. Mental and neurological disorders and risk of COVID-19 susceptibility, illness severity and mortality: A systematic review, meta-analysis and call for action. *EClinicalMedicine.* 2021;40:101111.
- Yuan K, Gong YM, Liu L, Sun YK, Tian SS, Wang YJ, et al. Prevalence of post-traumatic stress disorder after infectious disease pandemics in the twenty-first century, including COVID-19: a meta-analysis and systematic review. *Mol Psychiatry.* 2021;26:4982–98.
- Yuan K, Zheng YB, Wang YJ, Sun YK, Gong YM, Huang YT, et al. A systematic review and meta-analysis on prevalence of and risk factors associated with depression, anxiety and insomnia in infectious diseases, including COVID-19: a call to action. *Mol Psychiatry.* 2022;27:3214–22.
- Bao Y, Sun Y, Meng S, Shi J, Lu L. 2019-nCoV epidemic: address mental health care to empower society. *Lancet.* 2020;395:e37–e38.
- Gong Y, Liu X, Su S, Bao Y, Kosten TR, Lu L. Addressing mental health issues amid the COVID-19 pandemic: a wake-up call. *Sci Bull.* 2022;67:2259–62.
- Whitaker M, Elliott J, Chadeau-Hyam M, Riley S, Darzi A, Cooke G, et al. Persistent COVID-19 symptoms in a community study of 606,434 people in England. *Nat Commun.* 2022;13:1957.
- Fernández-de-Las-Peñas C, Martín-Guerrero JD, Florencio LL, Navarro-Pardo E, Rodríguez-Jiménez J, Torres-Macho J, et al. 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.* 2022;1–9. Epub ahead of print.
- China National Health Commission. Chinese clinical guidance for COVID-19 pneumonia diagnosis and treatment. <http://kjfy.meetingchina.org/msite/news/show/cn/3337.html>. Accessed 30 June 2022.
- Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest.* 1988;93:580–6.
- EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16:199–208.
- Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* 2005;53:695–9.
- Wassenaar A, de Reus J, Donders ART, Schoonhoven L, Cremer OL, de Lange DW, et al. Development and validation of an abbreviated questionnaire to easily measure cognitive failure in ICU survivors: a multicenter study. *Crit Care Med.* 2018;46:79–84.
- Heesakkers H, van der Hoeven JG, Corsten S, Janssen I, Ewalds E, Simons KS, et al. Clinical outcomes among patients with 1-Year survival following intensive care unit treatment for COVID-19. *JAMA.* 2022;327:559–65.
- Levis B, Benedetti A, Thombs B. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. *BMJ.* 2019;365:11781.
- Johnson SU, Ulvenes PG, Øktedalen T, Hoffart A. Psychometric properties of the General Anxiety Disorder 7-item (GAD-7) scale in a heterogeneous psychiatric sample. *Front Psychol.* 2019;10:1713.
- Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep.* 2011;34:601–8.
- Wang L, Zhang L, Armour C, Cao C, Qing Y, Zhang J, et al. Assessing the underlying dimensionality of DSM-5 PTSD symptoms in Chinese adolescents surviving the 2008 Wenchuan earthquake. *J Anxiety Disord.* 2015;31:90–97.
- Mplus User's Guide. <https://www.statmodel.com/html Ug.shtml>. Accessed 30 June 2022.
- Using Mplus TECH11 and TECH14 to test the number of latent classes. <https://www.statmodel.com/examples/webnotes/webnote14.pdf>. Accessed 30 June 2022.
- Rogers JP, Chesney E, Oliver D, Pollak TA, McGuire P, Fusar-Poli P, et al. Psychiatric and neuropsychiatric presentations associated with severe coronavirus infections: a systematic review and meta-analysis with comparison to the COVID-19 pandemic. *Lancet Psychiatry.* 2020;7:611–27.
- Taquet M, Sillett R, Zhu L, Mendel J, Camplisson I, Dercon Q, et al. Neurological and psychiatric risk trajectories after SARS-CoV-2 infection: an analysis of 2-year retrospective cohort studies including 1 284 437 patients. *Lancet Psychiatry.* 2022;9:815–27.
- Ali Awan H, Najmuddin Diwan M, Aamir A, Ali M, Di Giannantonio M, Ullah I, et al. SARS-CoV-2 and the brain: what do we know about the causality of 'cognitive COVID'? *J Clin Med.* 2021;10:3441.
- Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. Considering how biological sex impacts immune responses and COVID-19 outcomes. *Nat Rev Immunol.* 2020;20:442–7.
- Sylvester SV, Rusu R, Chan B, Bellows M, O'Keefe C, Nicholson S. Sex differences in sequelae from COVID-19 infection and in long COVID syndrome: a review. *Curr Med Res Opin.* 2022;38:1–9.
- Menges D, Ballouz T, Anagnostopoulos A, Aschmann HE, Domenghino A, Fehr JS, et al. Burden of post-COVID-19 syndrome and implications for healthcare service planning: A population-based cohort study. *PLoS One.* 2021;16:e0254523.
- Valiente De Santis L, Pérez-Camacho I, Sobrino B, González GE, Ruíz-Mesa JD, Plata A, et al. Clinical and immunoserological status 12 weeks after infection with COVID-19: prospective observational study. *Epidemiol Int J.* 2020;4:1–7. <https://doi.org/10.23880/eij-16000167>.
- Marshall M. The four most urgent questions about long COVID. *Nature.* 2021;594:168–70.
- Blomberg B, Mohn KG, Brokstad KA, Zhou F, Linchusen DW, Hansen BA, et al. Long COVID in a prospective cohort of home-isolated patients. *Nat Med.* 2021;27:1607–13.
- Komaroff AL, Bateman L. Will COVID-19 lead to myalgic encephalomyelitis/chronic fatigue syndrome? *Front Med.* 2020;7:606824.
- Lam MH, Wing YK, Yu MW, Leung CM, Ma RC, Kong AP, et al. Mental morbidities and chronic fatigue in severe acute respiratory syndrome survivors: long-term follow-up. *Arch Intern Med.* 2009;169:2142–7.
- Epstein L, Wong KK, Kallen AJ, Uyeki TM. Post-Ebola signs and symptoms in U.S. survivors. *N Engl J Med.* 2015;373:2484–6.
- Parker AM, Brigham E, Connolly B, McPeake J, Agranovich AV, Kenes MT, et al. Addressing the post-acute sequelae of SARS-CoV-2 infection: a multidisciplinary model of care. *Lancet Respir Med.* 2021;9:1328–41.
- Kress JP, Hall JB. ICU-acquired weakness and recovery from critical illness. *N Engl J Med.* 2014;371:287–8.
- Kim SY, Kumble S, Patel B, Pruski AD, Azola A, Tatini AL, et al. Managing the rehabilitation wave: rehabilitation services for COVID-19 survivors. *Arch Phys Med Rehabil.* 2020;101:2243–9.



## ACKNOWLEDGEMENTS

This study was supported by a grant from the National Key Research and Development Program of China (2021YFC0863700, 2019YFA0706200), National Programs for Brain Science and Brain-like Intelligence Technology of China (STI2030-Major Projects, 2021ZD0200800, 2021ZD0200700), and the National Natural Science Foundation of China (no. 81761128036, 82171514).

## AUTHOR CONTRIBUTIONS

LL and YB proposed the topic and main idea; YZ, LS, NZ, HM, SN, SW, SSu, YZ, KY, WY, SM, JSun, JShi, and YB designed the questionnaire, YZ, ZJ, YL, JY, FJ, XL, YB, and LL contributed to recruiting participants and the data collection. YZ, NZ analyzed the data and wrote the initial draft of the manuscript; YZ, LS, NZ, YB, ZK, DA, JS and LL commented on and revised the manuscript; LL, YB, and DA finalized the manuscript, with contributions from all authors. All authors have read and agreed to the published version of the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41380-023-01951-1>.

**Correspondence** and requests for materials should be addressed to Xiangyou Li, Yanping Bao or Lin Lu.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.