JAMA | Original Investigation

Development of a Definition of Postacute Sequelae of SARS-CoV-2 Infection

Tanayott Thaweethai, PhD; Sarah E. Jolley, MD, MS; Elizabeth W. Karlson, MD, MS; Emily B. Levitan, ScD; Bruce Levy, MD; Grace A. McComsey, MD; Lisa McCorkell, MPP; Girish N. Nadkarni, MD, MPH; Sairam Parthasarathy, MD; Upinder Singh, MD; Tiffany A. Walker, MD; Caitlin A. Selvaggi, MS; Daniel J. Shinnick, MS; Carolin C. M. Schulte, PhD; Rachel Atchley-Challenner, PhD; RECOVER Consortium Authors; Leora I. Horwitz, MD; Andrea S. Foulkes, ScD; for the RECOVER Consortium

IMPORTANCE SARS-CoV-2 infection is associated with persistent, relapsing, or new symptoms or other health effects occurring after acute infection, termed *postacute sequelae of SARS-CoV-2 infection* (PASC), also known as *long COVID*. Characterizing PASC requires analysis of prospectively and uniformly collected data from diverse uninfected and infected individuals.

OBJECTIVE To develop a definition of PASC using self-reported symptoms and describe PASC frequencies across cohorts, vaccination status, and number of infections.

DESIGN, SETTING, AND PARTICIPANTS Prospective observational cohort study of adults with and without SARS-CoV-2 infection at 85 enrolling sites (hospitals, health centers, community organizations) located in 33 states plus Washington, DC, and Puerto Rico. Participants who were enrolled in the RECOVER adult cohort before April 10, 2023, completed a symptom survey 6 months or more after acute symptom onset or test date. Selection included population-based, volunteer, and convenience sampling.

EXPOSURE SARS-CoV-2 infection.

MAIN OUTCOMES AND MEASURES PASC and 44 participant-reported symptoms (with severity thresholds).

RESULTS A total of 9764 participants (89% SARS-CoV-2 infected; 71% female; 16% Hispanic/Latino; 15% non-Hispanic Black; median age, 47 years [IQR, 35-60]) met selection criteria. Adjusted odds ratios were 1.5 or greater (infected vs uninfected participants) for 37 symptoms. Symptoms contributing to PASC score included postexertional malaise, fatigue, brain fog, dizziness, gastrointestinal symptoms, palpitations, changes in sexual desire or capacity, loss of or change in smell or taste, thirst, chronic cough, chest pain, and abnormal movements. Among 2231 participants first infected on or after December 1, 2021, and enrolled within 30 days of infection, 224 (10% [95% CI, 8.8%-11%]) were PASC positive at 6 months.

CONCLUSIONS AND RELEVANCE A definition of PASC was developed based on symptoms in a prospective cohort study. As a first step to providing a framework for other investigations, iterative refinement that further incorporates other clinical features is needed to support actionable definitions of PASC.

Editorial page 1918

Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Andrea S. Foulkes, ScD, Massachusetts General Hospital Biostatistics, 50 Staniford St, Ste 560, Boston, MA 02114 (afoulkes@mgh.harvard.edu).

JAMA. 2023;329(22):1934-1946. doi:10.1001/jama.2023.8823
Published online May 25, 2023.

More than 658 million people worldwide have been infected with SARS-CoV-2.[1] Postacute sequelae of SARS-CoV-2 infection (PASC), also known as *long COVID* and defined as ongoing, relapsing, or new symptoms or conditions present 30 or more days after infection, is a major clinical and public health concern.[2-6] Short- and long-term effects of PASC have substantial impacts on health-related quality of life, earnings, and health care costs.[7,8] Most existing PASC studies have focused on individual symptom frequency and have generated widely divergent estimates of prevalence due to their retrospective design and lack of an uninfected comparison group. Moreover, defining PASC precisely is difficult because it is heterogeneous, composed of conditions with variable and potentially overlapping etiologies (eg, organ injury, viral persistence, immune dysregulation, autoimmunity, and gut dysbiosis).[9,10]

It is of significant public health and scientific importance to better research the underlying mechanisms of PASC and potential preventive and therapeutic interventions. This effort requires data collection on SARS-CoV-2–infected and -uninfected individuals in a large prospective cohort study designed specifically to characterize PASC. Additionally, simultaneous consideration of multiple symptoms that persist over time and application of appropriate analytical techniques are essential. Further consideration of changes in PASC frequency and its manifestations over the course of the COVID-19 pandemic, due to variable SARS-CoV-2 strains, new treatment and prevention strategies, and repeat infections, is important.

This study is part of the National Institutes of Health's Researching COVID to Enhance Recovery (RECOVER) Initiative, which seeks to understand, treat, and prevent PASC (https://recovercovid.org/). In this first analysis of data from the RECOVER adult cohort, criteria for identifying PASC based on self-reported symptoms are delineated and several distinctive PASC subphenotypes with varying impacts on well-being and physical health are described. This study was enriched with self-referred participants to promote inclusive participation. Estimates were expected to be more accurate in the subcohort of participants enrolled within 30 days of acute infection, for whom selection bias based on PASC would be minimal.

Unlike electronic health records and most existing cohort studies, data from this study captured PASC-specific self-reported symptoms based on standardized questionnaires developed with input from patient representatives. This report is an adequately powered, prospective study of PASC based on participant-reported symptoms that included both infected and uninfected individuals over the course of the pandemic. Notably, unlike prior reports, the paradigm presented here does not rely on predefined clinical symptoms; instead, a definition of PASC as a new condition specific to SARS-CoV-2 infection is proposed.

Key Points

Question What symptoms are differentially present in SARS-CoV-2–infected individuals 6 months or more after infection compared with uninfected individuals, and what symptom-based criteria can be used to identify postacute sequelae of SARS-CoV-2 infection (PASC) cases?

Findings In this analysis of data from 9764 participants in the RECOVER adult cohort, a prospective longitudinal cohort study, 37 symptoms across multiple pathophysiological domains were identified as present more often in SARS-CoV-2–infected participants at 6 months or more after infection compared with uninfected participants. A preliminary rule for identifying PASC was derived based on a composite symptom score.

Meaning A framework for identifying PASC cases based on symptoms is a first step to defining PASC as a new condition. These findings require iterative refinement that further incorporates clinical features to arrive at actionable definitions of PASC.

Methods

Institutional review boards at NYU Grossman School of Medicine, serving as a single institutional review board, and other participating institutions reviewed and approved the protocol. All participants provided written informed consent to participate in research.

Study Design

The RECOVER adult cohort study included SARS-CoV-2–infected and -uninfected participants (the trial protocol is in Supplement 1 and the statistical analysis plan in Supplement 2). All infected participants met World Health Organization suspected, probable, or confirmed criteria.[11] Index for infected participants was defined as date of first positive SARS-CoV-2 test result or COVID-19 symptom onset. Uninfected participants had no known history of SARS-CoV-2 infection and index was defined as a past negative SARS-CoV-2 test result date. Participants belonged to either the acute cohort (enrolled ≤30 days since index) or the postacute cohort (enrolled >30 days to 3 years after index). Participants were recruited from 85 sites across the United States and completed office visits and remote surveys developed with early engagement of patients, support group stakeholders, and multidisciplinary clinical experts.[12]

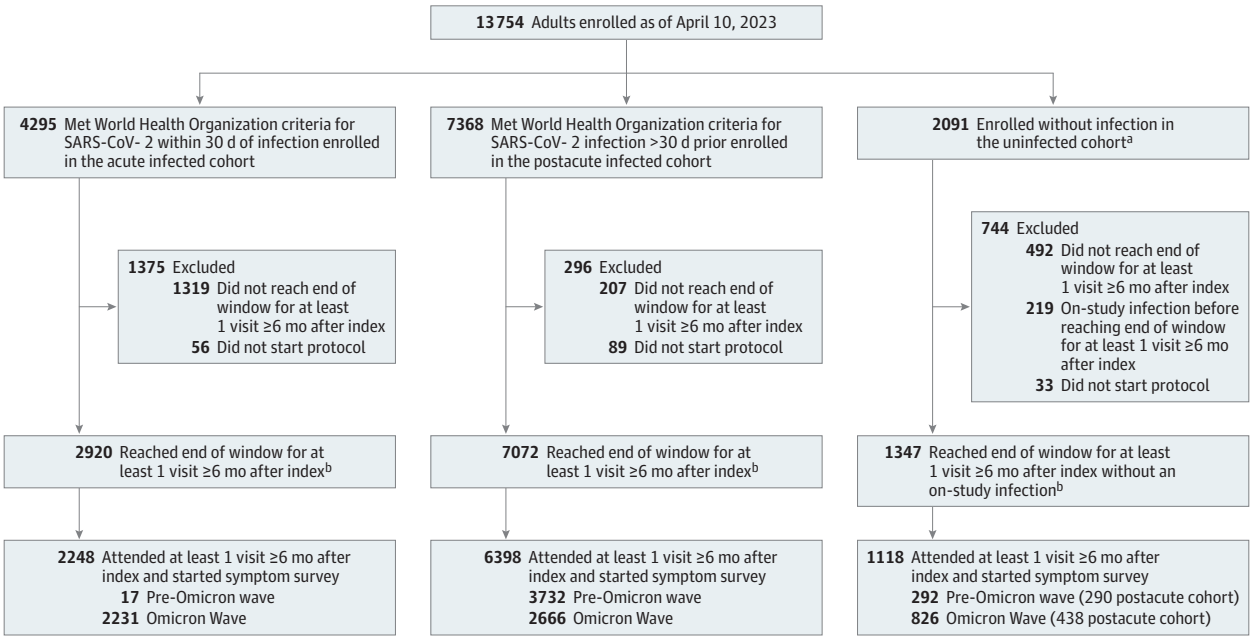
Participants

Adult participants enrolled prior to April 10, 2023 (N = 13 754) were considered (**Figure 1**). Enrollment is ongoing, and not all enrolled participants have reached eligibility for inclusion. The analysis cohort included participants with a study visit completed 6 months or more after the index date (**Table 1**). Uninfected participants with a reported on-study infection and participants who had no symptom survey data were excluded. A subgroup of participants also belonged to the RECOVER pregnancy cohort. Race and ethnicity were captured via participant self-report using fixed categories to better understand racial and ethnic differences in sequelae due to SARS-CoV-2 infection (eMethods in Supplement 3).

Exposures and Outcomes

The analysis used the first study visit at 6 months or more after the index date. The exposure was SARS-CoV-2 infection prior

Figure 1. RECOVER Adult Analysis Cohort



[a] Uninfected included acute cohort (n = 1092) and postacute cohort (n = 999) participants. Uninfected participants had no known history of SARS-CoV-2 infection. Acute uninfected participants were enrolled within 30 days of a SARS-CoV-2 negative test result, while postacute uninfected participants were enrolled more than 30 days after a SARS-CoV-2 negative test result.

[b] Participants who completed visit without reaching end of visit window were included in this count.

to study enrollment. Uninfected participants with antibody results at enrollment indicating prior infection were reclassified as infected and assigned an index date 90 days prior. The primary outcome was the presence of each of 44 symptoms (eTable 1 in Supplement 3). Using these symptoms, a PASC definition was developed. The primary analysis used symptom presence for inclusivity; sensitivity analysis considered new-onset symptoms. Results were reported for 3 additional age- and sex-dependent symptoms (eTable 1 in Supplement 3).

Statistical Analysis

Results were reported overall and within 3 subcohorts: acute Omicron (n = 2231 infected, n = 388 uninfected; index date on or after December 1, 2021); postacute pre-Omicron (n = 3732 infected, n = 290 uninfected; index date before December 1, 2021); and postacute Omicron (n = 2666 infected, n = 438 uninfected; index date on or after December 1, 2021). Acute cohort participants with a pre-Omicron index date (17 infected, 2 uninfected) were included in overall analyses.

Balancing weights were used to account for differences in the age, sex, and race and ethnicity distributions between infected and uninfected participants (eMethods in Supplement 3). Symptom frequency was defined as the proportion reporting a symptom and exceeding corresponding moderate to severe symptom severity threshold (eTables 1 and 2 in Supplement 3). Symptoms with frequency of 2.5% or greater were considered. Symptom frequencies by infection status were reported and adjusted odds ratios (aORs) were calculated using weighted logistic regression. In sensitivity analysis, new-onset symptom frequency was defined as the proportion of participants with the symptom at study visit among those without the symptom in the year prior to the index date. Symptom frequencies characterized the study cohort and were not unbiased estimators of population-level prevalence due to the cohort sampling strategies. Symptom frequency estimates within the acute Omicron subcohort were expected to be more aligned with the corresponding population frequencies.

A rule for identifying PASC was derived. Symptoms differentiating infected and uninfected participants were identified using least absolute shrinkage and selection operator (LASSO) with balancing weights.[13] Each symptom was assigned a score based on the estimated coefficients and participants were assigned a total score by summing the symptom scores for each reported symptom. Using 10-fold cross-validation, an optimal score threshold for PASC was selected (eMethods in Supplement 3). Participants meeting the PASC score threshold were classified as PASC positive; others were classified as PASC unspecified. The proportions were reported.

Participants classified as PASC positive were clustered into subgroups using unsupervised learning (K-means consensus clustering[14] followed by hierarchical clustering[15]) including symptoms identified with LASSO. Symptoms highly correlated with those identified by LASSO were reported. The distribution of PASC score and Patient-Reported Outcomes Measurement Information System (PROMIS) Global Health 10

Table 1. RECOVER Adult Cohort Demographic Characteristics by Infection Status at Enrollment

Characteristic[a]	No./total (%)		
	Infected (n = 8646)	Uninfected (n = 1118)	Uninfected with balancing weights (n = 1109), %
Age at enrollment, y			
Median (IQR)	45 (34-59)	55 (40-65)	45 (35-60)
No.	8637	1117	1109
Age category at enrollment, y			
18-45	4389/8637 (51)	377/1117 (34)	51
46-65	3175/8637 (37)	502/1117 (45)	37
>65	1073/8637 (12)	238/1117 (21)	12
Race and ethnicity[b]			
Asian, non-Hispanic	428/8558 (5)	73/1106 (7)	5
Black or African American, non-Hispanic	1220/8558 (14)	197/1106 (18)	14
Hispanic, Latino, or Spanish	1473/8558 (17)	119/1106 (11)	17
White, non-Hispanic	5027/8558 (59)	685/1106 (62)	59
Multiracial/multiethnic	305/8558 (4)	26/1106 (2)	4
Other	105/8558 (1)	6/1106 (1)	1
Sex assigned at birth			
Female	6221/8602 (72)	711/1110 (64)	72
Male	2377/8602 (28)	399/1110 (36)	28
Intersex	4/8602 (<1)	0/1110	0
Vaccination status at index date[c]			
Unvaccinated	3291/8538 (39)	161/1095 (15)	16
Partially vaccinated	154/8538 (2)	21/1095 (2)	2
Fully vaccinated	4725/8538 (55)	860/1095 (79)	77
Date of last dose unknown	368/8538 (4)	53/1095 (5)	5
Cohort and prevalent SARS-CoV-2 strain at Index[d]			
Acute pre-Omicron	17/8646 (<1)	2/1118 (<1)	<1
Acute Omicron	2231/8646 (26)	388/1118 (35)	33
Postacute pre-Omicron	3732/8646 (43)	290/1118 (26)	28
Postacute Omicron	2666/8646 (31)	438/1118 (39)	39
Medically underserved area			
Yes	2369/8646 (27)	298/1118 (27)	28
No	6277/8646 (73)	820/1118 (73)	72
Rural participant			
Yes	465/8646 (5)	45/1118 (4)	4
No	8181/8646 (95)	1073/1118 (96)	96
Education level			
Did not complete high school/no diploma	280/8564 (3)	44/1107 (4)	3
High school/GED/some college/vocational/technical	2790/8564 (33)	306/1107 (28)	26
Bachelor's/advanced degree	5494/8564 (64)	757/1107 (68)	70

[a] Additional characteristics were summarized in eTable 5 in Supplement 3. Summaries by subcohort were provided in eTable 6 in Supplement 3.

[b] Race and ethnicity were captured via participant self-report using fixed categories to better understand racial and ethnic differences in sequelae due to SARS-CoV-2 infection (eMethods in Supplement 3).

[c] Details on how vaccination status was defined were provided in eTable 4 in Supplement 3.

[d] The pre-Omicron prevalent strain was defined as prior to December 1, 2021.

general quality of life (Q2), general physical health (Q3), and ability to carry out everyday physical activities (Q6) (eTable 3 in Supplement 3) were reported.

Rates of PASC were assessed by infection status, sex, age, and vaccination status at the index date (eTable 4 in Supplement 3), reinfection (between index and analysis visit), and visit month. The proportion of participants meeting criteria for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS, defined based on RECOVER survey; eMethods in Supplement 3) who were PASC positive at the same visit was reported. Sensitivity analyses removed the symptom severity thresholds and separately added well-being and physical health requirements for PASC (Q2 or Q3: fair or poor; Q6: moderate or worse). Inverse probability weighting was

Table 2. Model-Selected Symptoms That Define PASC and Their Corresponding Scores[a]

Symptom	Log odds ratio	Score
Smell/taste	0.776	8
Postexertional malaise	0.674	7
Chronic cough	0.438	4
Brain fog[b]	0.325	3
Thirst	0.255	3
Palpitations	0.238	2
Chest pain[b]	0.233	2
Fatigue[b]	0.148	1
Sexual desire or capacity	0.126	1
Dizzines	0.121	1
Gastrointestinal	0.085	1
Abnormal movements	0.072	1
Hair loss	0.049	0

Abbreviation: PASC, postacute sequelae of SARS-CoV-2 infection.

[a] Least absolute shrinkage and selection operator was used to identify which symptoms defined PASC. A symptom score was assigned by dividing the estimated log odds ratio by 0.10 and rounding to the nearest integer. For each person, the total score was defined as the sum of the scores for each symptom a person reported.

[b] Additional severity criteria required (eTables 1 and 2 in Supplement 3).

applied to account for loss to follow-up in the acute Omicron subcohort (eMethods in Supplement 3).

Results

Participants

A total of 9764 participants (8646 infected; 1118 uninfected) met study criteria (Figure 1, 71% female [6932/9712]; 16% Hispanic/Latino [1592/9664]; 15% non-Hispanic Black [1417/9664]; 58% fully vaccinated at the index date [5585/9633]; median age, 47 years [IQR, 35-60]). After application of balancing weights, the distributions of age, sex, and race and ethnicity were the same in infected and uninfected participants (Table 1). In the weighted cohort, uninfected participants were more likely to be fully vaccinated (77% vs 55%). Comorbidity frequencies were similar between infected and uninfected participants (eTable 5 in Supplement 3). Uninfected participants were more likely to be self-referrals or recruited via community outreach (eTable 5 in Supplement 3). A total of 1260 of 6932 female participants (18%) were in the pregnancy cohort.

Symptom Frequency

In the full cohort, 37 symptoms had frequency of 2.5% or greater and aORs were 1.5 or greater (infected vs uninfected participants) for all 37 (eFigure 1 in Supplement 3). Symptoms (using severity thresholds) with more than 15% absolute difference in frequencies (infected vs uninfected) included postexertional malaise (PEM) (28% vs 7%; aOR, 5.2 [95% CI, 3.9-6.8]), fatigue (38% vs 17%; aOR, 2.9 [95% CI, 2.4-3.4]), dizziness (23% vs 7%; aOR, 3.4 [95% CI, 2.6-4.4]), brain fog (20% vs 4%; aOR, 4.5 [95% CI, 3.2-6.2]), and gastrointestinal (GI) symptoms (25% vs 10%; aOR, 2.7 [95% CI, 2.2-3.4]).

Symptom-Level Sensitivity Analyses

In infected participants, the frequencies of new-onset symptoms (with severity thresholds) were similar, including PEM (28%), fatigue (37%), dizziness (21%), brain fog (20%), and GI symptoms (20%) (eFigure 2 in Supplement 3). The corresponding observed symptom frequencies without severity thresholds were higher (eg, fatigue, 47%; brain fog, 40%) (eFigure 3 in Supplement 3).

Symptom-Level Analysis by Subcohort

The distributions of demographics and comorbidities were comparable across the acute Omicron, postacute pre-Omicron, and postacute Omicron subcohorts, though there was a higher proportion unvaccinated in the postacute pre-Omicron subcohort (eTable 6 in Supplement 3). Time from the index date to analysis visit ranged from 6 to 15 months in the acute Omicron and postacute Omicron subcohorts and 6 to 39 months in the postacute pre-Omicron subcohort (eFigure 4 in Supplement 3). Generally, symptom frequencies and the differences between infected and uninfected participants were lower in the acute Omicron subcohort, higher in the postacute Omicron subcohort, and highest in the postacute pre-Omicron subcohort (eFigures 5-7 in Supplement 3). Symptom frequencies in acute Omicron participants who were also fully vaccinated were the lowest (eFigure 8 in Supplement 3).

PASC Score

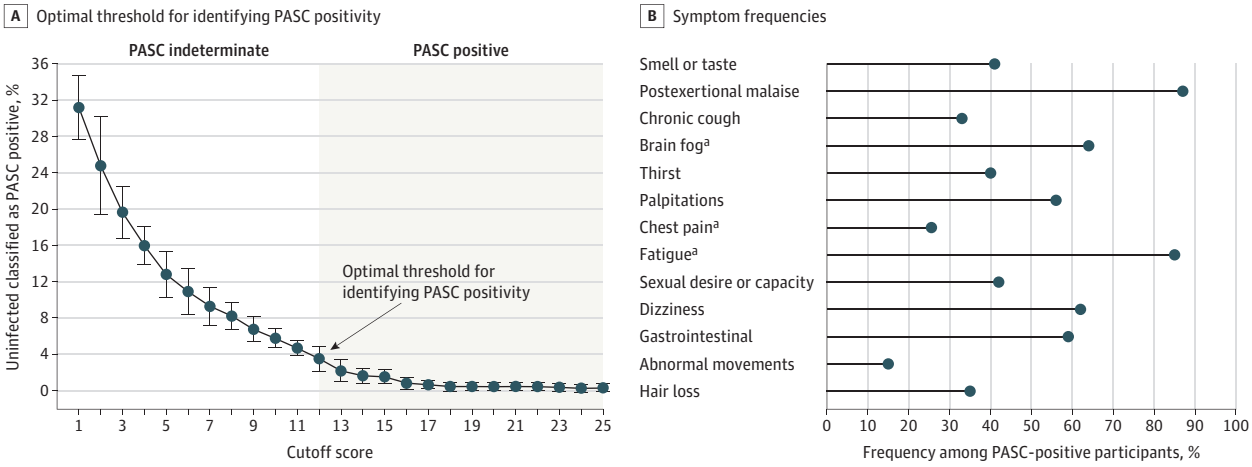
Using the full cohort, LASSO identified 12 symptoms with corresponding scores ranging from 1 to 8 (**Table 2**). The optimal PASC score threshold used was 12 or greater (**Figure 2**A). The symptoms (ordered by decreasing frequencies among participants with a qualifying PASC score) were PEM, fatigue, brain fog, dizziness, GI symptoms, palpitations, changes in sexual desire or capacity, loss of or change in smell or taste, thirst, chronic cough, chest pain, and abnormal movements. Symptoms correlated with the selected symptoms included dry mouth, weakness, headaches, tremor, muscle and abdominal pain, fever/sweats/chills, and sleep disturbance (eTable 7 in Supplement 3).

The proportion with a qualifying PASC score in the full cohort (subject to selection bias) was 1990 of 8646 infected participants (23%) and 41 of 1118 of uninfected participants (3.7%) (overall: 2031/9764 [21%]). Among participants with PASC, the most common symptoms were PEM (87%), fatigue (85%), brain fog (64%), dizziness (62%), GI (59%), and palpitations (57%) (Figure 2B; eTable 8 in Supplement 3). Higher PASC scores were associated with worse PROMIS Global 10 scores (Figure 2C).

PASC in Subcohorts

The proportion of infected participants with PASC in the acute Omicron subcohort was 10% (95% CI, 8.8%-11%; 224/2231). After adjustment for missing data, the estimated rate was 9.8% (95% CI, 8.6%-11%). It was greater in the postacute pre-Omicron (1320/3732 [35%; 95% CI, 34%-37%]) and postacute Omicron (442/2666 [17%; 95% CI, 15%-18%]) subcohorts (**Table 3**). Symptom frequencies among PASC-infected participants were similar across subcohorts, with a few notable exceptions, including brain fog, GI symptoms, and palpitations

Figure 2. Defining the Postacute Sequelae of SARS-CoV-2 Infection (PASC) Score and a Decision Rule



C Distribution of PROMIS Global 10 responses

PROMIS Global 10 Q2: general quality of life

	0	1-2	3-6	7-11	12-16	≥17
Excellent	29	17	13	12	6	4
Very good	44	46	38	36	22	17
Good	21	30	37	36	41	32
Fair	5	6	10	14	25	33
Poor	1	0	2	2	7	14
No. of participants	3951	1412	1106	1264	998	1033

PASC score (quintile above 0), %

PROMIS Global 10 Q3: general physical health

	0	1-2	3-6	7-11	12-16	≥17
Excellent	15	6	4	4	1	0
Very good	41	31	27	20	11	7
Good	31	43	42	41	34	26
Fair	10	18	24	29	42	41
Poor	1	1	4	6	12	26
No. of participants	3951	1412	1106	1264	998	1033

PASC score (quintile above 0), %

PROMIS Global 10 Q6: ability to carry out everyday physical activities

	0	1-2	3-6	7-11	12-16	≥17
Completely	78	71	59	47	27	15
Mostly	12	17	21	25	27	22
Moderately	6	9	14	20	28	30
A little	3	3	6	8	17	31
Not at all	0	0	0	0	1	1
No. of participants	3951	1412	1106	1264	998	1033

PASC score (quintile above 0), %

A, Optimal score cutoff for classifying a participant as PASC positive using cross-validation (eMethods in Supplement 3). The decision rule based on symptoms is intended to identify participants with PASC. PASC status for participants not meeting the score threshold requires consideration of additional data inputs.

B, Symptom frequencies among PASC-positive participants for symptoms that contribute to the PASC score. Many other symptoms have high frequency in PASC-positive participants (eTable 8 in Supplement 3).

C, Distribution of Patient-Reported Outcomes Measurement Information System (PROMIS) Global 10 responses among participants with a zero PASC score and among participants within nonzero PASC score quintiles. The PROMIS Global 10 provides an assessment of quality of life along 10 dimensions, each rated on a 5-point scale. The shading corresponds to frequency within each column on a scale from 0% to 100%.

[a] Additional severity criteria required (eTables 1 and 2 in Supplement 3).

(eTable 8 in Supplement 3). The proportion of PASC positivity was lower among fully vaccinated than unvaccinated participants (acute Omicron: 9.7% vs 17%; postacute pre-Omicron: 31% vs 37%; postacute Omicron: 16% vs 22%) (Table 3). In the Omicron cohorts, the estimated proportion of PASC positivity was greater among reinfected participants compared with participants with 1 reported infection (acute Omicron: 20% vs 9.7%; postacute Omicron: 21% vs 16%) (Table 3).

PASC-Stratified Analysis in the Full Cohort

Among infected participants in the full cohort, the proportions of PASC positivity were 39% (299/757) among hospitalized participants and 22% (1636/7387) among not hospitalized participants during acute infection; 19% (442/2377) among males and 25% (1540/6221) among females; and 20% (885/4389) among those aged 18 to 45 years and 28% (904/3175) (28%) among those aged 46 to 65 years. In cross-sectional analysis, the proportion of PASC-positive participants was consistent over the visit month used in analysis (eTable 9 in Supplement 3). In subgroups with repeated visits at 6 and 9 months after the index date, PASC positivity varied over study visit, though 68% of PASC-positive participants remained positive at the subsequent visit (eTable 10 in Supplement 3). Among infected participants meeting criteria for ME/CFS, 98% met the criteria for PASC.

PASC Sensitivity Analyses in the Full Cohort

The proportions of PASC increased to 27% (2300/8646) among infected participants and 4.7% (52/1118) among uninfected participants when severity scores were not included in the PASC score determination. After applying additional PROMIS Global 10 criteria to qualify as PASC positive, 17% (1434/8646) were PASC positive among infected participants and 3.0% (34/1118) among uninfected participants.

PASC Subgroups

Four PASC subgroups were identified (**Figure 3**A). Features of PASC subgroups included loss of or change in smell or taste (100%) in cluster 1 (n = 477); PEM (99%) and fatigue (84%) in cluster 2 (n = 405); brain fog (100%), PEM (99%), and fatigue

Table 3. PASC Frequencies, Overall and Stratified by Subcohort, Infection, Vaccination, and Reinfection

	Total No.	PASC positive, No. (%)
PASC frequencies overall and stratified by subcohort and infection[a]		
All participants (full cohort)[a]	9764	2031 (21)
Infected	8646	1990 (23)
Uninfected	1118	41 (3.7)
Acute Omicron		
Infected	2231	224 (10)
Uninfected	388	18 (4.6)
Postacute pre-Omicron		
Infected	3732	1320 (35)
Uninfected	290	11 (3.8)
Postacute Omicron		
Infected	2666	442 (17)
Uninfected	438	12 (2.7)
PASC frequencies stratified by study cohort, prevalent SARS-CoV-2 strain, and vaccination (infected participants only)[b]		
Acute Omicron		
Vaccinated	2016	195 (9.7)
Not vaccinated	86	15 (17)
Postacute pre-Omicron		
Vaccinated	491	154 (31)
Not vaccinated	2967	1090 (37)
Postacute Omicron		
Vaccinated	2208	356 (16)
Not vaccinated	232	50 (22)
PASC frequencies stratified by study cohort, prevalent SARS-CoV-2 strain, and reinfection (infected participants only)		
Acute Omicron		
1 infection	2150	208 (9.7)
>1 infections	81	16 (20)
Postacute pre-Omicron		
1 infection	2484	907 (37)
>1 infections	1248	413 (33)
Postacute Omicron		
1 infection	2200	345 (16)
>1 infections	466	97 (21)

Abbreviation: PASC, postacute sequelae of SARS-CoV-2 infection.

[a] Acute cohort participants with a pre-Omicron index date were included in the full cohort analysis.

[b] Participants who were partially vaccinated, missing vaccination status, or missing date of last dose (eTable 4 in Supplement 3) were excluded.

(94%) in cluster 3 (n = 587); and fatigue (94%), PEM (94%), dizziness (94%), brain fog (94%), GI (88%), and palpitations (86%) in cluster 4 (n = 562) (Figure 3B). Twenty-six percent of the PASC-unspecified group also met PROMIS Global 10 criteria compared with 53% of participants in cluster 1, 69% in cluster 2, 77% in cluster 3, and 86% in cluster 4. Among infected participants, 456 of 1540 females (30%) and 102 of 442 males (23%) with PASC were in cluster 4. A total of 277 of 885 participants aged 18 to 45 years (31%), 254 of 904 participants aged 46 to 65 years (28%), and 29 of 198 participants aged older than 65 years (15%) with PASC were in cluster 4.

PASC Subgroups by Subcohort

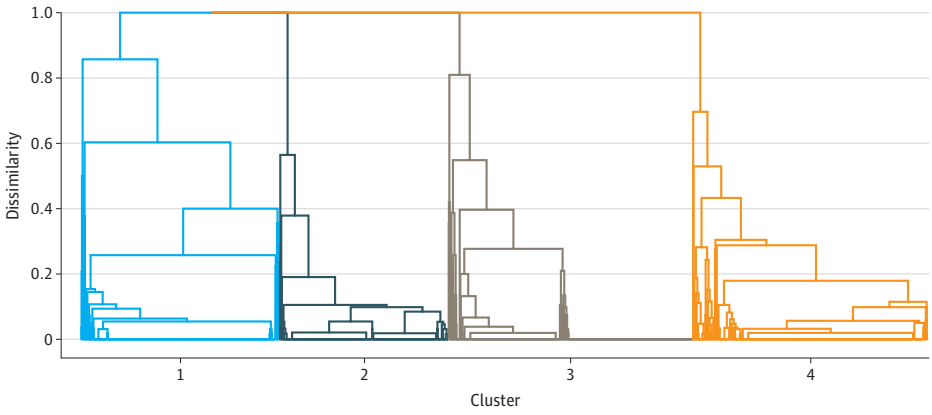
The proportion of PASC-positive infected participants in cluster 4 was higher within the postacute pre-Omicron (31%) than postacute Omicron (23%) and acute Omicron (23%) subcohorts (eTable 11A in Supplement 3). Overall, among PASC-positive infected participants, the proportion in cluster 4 among fully vaccinated compared with unvaccinated participants was 23% vs 32% (eTable 11B in Supplement 3). The distribution of clusters was similar for participants with a single reported infection compared with those with more than 1 infection, though the results varied by prevalent SARS-CoV-2 strain (eTable 11C in Supplement 3).

Discussion

This study reported early results from a prospective, survey-based cohort of adult SARS-CoV-2–infected and –uninfected individuals with ascertainment of patient-reported symptoms. A data-driven scoring framework was developed to

Figure 3. Identification of Postacute Sequelae of SARS-CoV-2 Infection (PASC) Subgroups and Their Characteristics

A Symptom profile clusters



B Symptom frequencies by PASC status

	Cluster 1 (n=477)	Cluster 2 (n=405)	Cluster 3 (n=587)	Cluster 4 (n=562)	PASC status: Positive (n=2031)	PASC status: Unspecified (n=7733)
Fatigue[a]	66	84	94	94	85	23
Fever, sweats, or chills	27	26	28	56	35	5
Postexertional malaise[a]	55	99	99	94	87	9
Swelling of legs	22	23	18	40	26	6
Chest pain[a]	13	26	14	50	26	2
Palpitations[a]	38	59	44	86	57	9
Hair loss[a]	29	31	21	59	35	13
Skin color changes	10	17	16	38	21	3
Skin pain	5	6	8	18	9	1
Skin rash	12	16	17	32	20	5
Hearing	40	38	40	62	46	14
Vision	19	20	25	51	30	4
Abdominal pain	7	11	12	36	17	2
Dry mouth	29	42	26	55	38	7
Gastrointestinal[a]	42	60	45	88	59	14
Teeth	18	21	17	43	25	8
Thirst[a]	30	48	20	62	40	6
Back pain	26	31	32	58	38	8
Foot pain	15	19	15	36	22	3
Joint pain	32	33	36	64	42	9
Muscle pain	27	33	34	60	39	6
Weakness	24	33	41	67	42	4
Abnormal movements[a]	5	10	8	33	15	1
Brain fog[a]	38	0	100	94	64	7
Dizziness[a]	31	56	62	94	62	10
Headache	23	26	37	64	39	5
Smell or taste[a]	100	3	6	53	41	4
Tremor	14	14	15	34	20	3
Anxiety	18	17	30	40	27	7
Depression	17	17	32	44	29	6
Sexual desire or capacity[a]	29	33	35	66	42	11
Sleep disturbance	18	22	32	49	32	5
Chronic cough[a]	33	43	16	43	33	5
Shortness of breath	22	30	31	58	36	3
Sleep apnea	30	36	32	44	36	12
Throat pain	5	5	7	24	11	1
Bladder	20	25	23	49	30	8

Cluster

PASC status

A, Dendrogram illustrating how PASC-positive participants with similar symptom profiles cluster. Each branch in the dendrogram represents a participant, and each cluster represents a subgroup of participants.

B. Heatmap of symptom frequencies within PASC unspecified and within each PASC-positive subgroup. The shading corresponds to frequency within each column on a scale from 0% to 100%.

[a] Although unsupervised learning uses 12 symptoms selected by least absolute shrinkage and selection operator (LASSO) (Figure 2), many other symptoms occur in combination with these 12.

classify PASC as a condition specific to SARS-CoV-2 infection. Based on this PASC score, 10% of participants first infected on or after December 1, 2021, and enrolled within 30 days of infection were classified as PASC positive at 6 months after infection. Increasing levels of the PASC score were associated with progressively worse measures of well-being and functioning. Although only 12 symptoms contributed to the PASC score, other symptoms correlated with this subgroup are individually important, considering their potential adverse impact on health-related quality of life.

PASC positivity was more common and associated with more severe manifestation in participants infected in the pre-Omicron era. Though participants with earlier infection may have been more likely to enroll in the RECOVER adult cohort because of known PASC, multiple studies reported an association between PASC and early pandemic variants.[16] Among participants with a first infection during the Omicron era, PASC frequency was higher among those with recurrent infections, corroborating electronic health record–based studies.[4,17-23] Though studies on the effect of vaccination are conflicting, these findings of modest reduction in PASC frequency among fully vaccinated participants align with recent systematic reviews.[24,25]

This study found that long-term symptoms associated with SARS-CoV-2 infection spanned multiple organ systems. The diversity of symptoms may be related to persistent viral reservoirs, autoimmunity, or direct differential organ injury. The symptoms identified are consistent with those reported in studies that assessed PASC manifestations (eMethods in Supplement 3). However, by simultaneously considering the contributions of multiple self-reported symptoms, a PASC-scoring algorithm that provides a framework for diagnosing PASC was developed.

Given the heterogeneity of PASC symptoms, determining whether PASC represents one unified condition or reflects a group of unique phenotypes is important. Recent evidence supports the presence of PASC phenotypes, although characterization of these phenotypes is inconsistent and largely dependent on available data.[2,6,22,23] Accurate phenotypic stratification has important implications for investigations into the pathophysiological processes underlying PASC and clinical trial design. PASC subgroups that demonstrate overlap with conditions previously described in clinical practice are detailed here, including olfactory dysfunction, cardiopulmonary sequelae, neurocognitive impairment, ME/CFS, and dysautonomia[26-30] and overlap with those reported by the National COVID Cohort Collaborative.[6] Biological samples from these participants may enable the development of biomarkers of PASC and reveal insights into the mechanistic underpinnings of PASC that inform choice of therapeutic interventions and case selection in upcoming clinical trials for PASC.

Limitations

First, the proposed paradigm and accompanying decision rule require iterative refinement as additional data become available. The PASC score provides an operational definition of PASC and requires further refinement and validation. RECOVER recruitment is ongoing, and not all participants have reached the analysis stage. Evolution and refinement of the phenotypes are anticipated as additional data become available.

Second, selection bias was likely among postacute cohort participants that may have affected frequency estimates including distribution of subphenotypes because PASC severity may impact study participation. Differential attrition of symptomatic and asymptomatic participants at follow-up visits could also have biased frequency estimates though use of inverse probability weighting in the acute cohorts mitigated this bias.

Third, uninfected participants may have had prior asymptomatic SARS-CoV-2 infections not detected due to variations in antibody production and persistence, weakening the discriminant characteristics of this PASC score threshold.

Fourth, symptoms were self-reported and only some symptoms integrate severity scales. Participants could report other symptoms as free text; these were not included in this analysis.

Fifth, confounding may have impacted effect sizes, eg, vaccination status may have been higher in participants at higher risk of PASC, attenuating a vaccination effect. Additionally, PASC status can change over time, perhaps due to underlying mechanistic changes.

Sixth, more than 200 symptoms of PASC have been reported, each with the potential of being life-altering and debilitating, and the symptoms highlighted herein may not reflect the severity or impact of other symptoms.

Conclusions

This symptom-based PASC definition represents a first step for identifying PASC cases and serves as a launching point for further investigations. Definition of a classification rule for PASC requires an updated algorithm that incorporates symptoms as well as biological features. Future analyses must consider the relationships among age, sex, race and ethnicity, social determinants of health, vaccination status after index date, comorbidities, and pregnancy status during infection on the risk of PASC and the distribution of PASC subgroups.

ARTICLE INFORMATION

Accepted for Publication: May 1, 2023.

Published Online: May 25, 2023. doi:10.1001/jama.2023.8823

Author Affiliations: Massachusetts General Hospital, Boston (Thaweethai, Selvaggi, Shinnick, Schulte, Foulkes); Harvard Medical School, Boston, Massachusetts (Thaweethai, Levy, Foulkes); University of Colorado Anschutz Medical Campus, Aurora (Jolley); Brigham and Women's Hospital, Boston, Massachusetts (Karlson, Levy); University of Alabama at Birmingham (Levitan); Case Western Reserve University, Cleveland, Ohio (McComsey); Patient-Led Research Collaborative, Calabasas, California (McCorkell); Icahn School of Medicine at Mount Sinai, New York, New York (Nadkarni); The University of Arizona College of Medicine, Tucson (Parthasarathy); Stanford University School of Medicine, Stanford, California (Singh); Emory University School of Medicine, Atlanta, Georgia (Walker); Mass General Brigham, Boston, Massachusetts (Atchley-Challenner); New York University Grossman School of Medicine, New York (Horwitz).

RECOVER Consortium Authors: George A. Alba, MD; Radica Alicic, MD; Natasha Altman, MD; Khamal Anglin, MD, MPH; Urania Argueta, BS; Hassan Ashktorab, PhD; Gaston Baslet, MD; Ingrid V. Bassett, MD, MPH; Lucinda Bateman, MD;

Brahmchetna Bedi, PhD; Shamik Bhattacharyya, MD, MS; Marie-Abele Bind, PhD; Andra L. Blomkalns, MD, MBA; Hector Bonilla, MD; Patricia A. Bush, MS, EdD; Mario Castro, MD, MPH; James Chan, MA; Alexander W. Charney, MD, PhD; Peter Chen, MD; Lori B. Chibnik, PhD, MPH; Helen Y. Chu, MD, MPH; Rebecca G. Clifton, PhD; Maged M. Costantine, MD; Sushma K. Cribbs, MD, MSc; Sylvia I. Davila Nieves, MS; Steven G. Deeks, MD; Alexandria Duven, RN; Ivette F. Emery, PhD; Nathan Erdmann, MD, PhD; Kristine M. Erlandson, MD, MS; Kacey C. Ernst, PhD, MPH; Rachael Farah-Abraham, PhD; Cheryl E. Farner, MSN; Elen M. Feuerriegel, PhD; Judes Fleurimont, MPH; Vivian Fonseca, MD; Nicholas Franko, BS; Vivian Gainer, MS; Jennifer C. Gander, PhD; Edward M. Gardner, MD; Linda N. Geng, MD, PhD; Kelly S. Gibson, MD; Minjoung Go, MD, MPH; Jason D. Goldman, MD, MPH; Halle Grebe, BS; Frank L. Greenway, MD; Mounira Habli, MD; John Hafner, MD, MPH; Jenny E. Han, MD, MS; Keith A. Hanson, MD, PhD; James Heath, PhD; Carla Hernandez, RN; Rachel Hess, MD, MS; Sally L. Hodder, MD; Matthew K. Hoffman, MD, MPH; Susan E. Hoover, MD, PhD; Beatrice Huang, BA; Brenna L. Hughes, MD; Prasanna Jagannathan, MD; Janice John, MS, MHCDS; Michael R. Jordan, MD; Stuart D. Katz, MD, MS; Elizabeth S. Kaufman, MD; John D. Kelly, MD; Sara W. Kelly, PhD, MPH; Megan M. Kemp, BA; John P. Kirwan, PhD; Jonathan D. Klein, MD, MPH; Kenneth S. Knox, MD; Jerry A. Krishnan, MD, PhD; Andre Kumar, MD; Adeyinka O. Laiyemo, MD; Allison A. Lambert, MD; Margaret Lanca, PhD; Joyce K. Lee-Iannotti, MD; Brian P. Logarbo, MD, MS; Michele T. Longo, MD; Carlos A. Luciano, MD; Karen Lutrick, PhD; Jason H. Maley, MD, MS; Jai G. Marathe, MD, MBBS; Vincent Marconi, MD; Gailen D. Marshall, MD, PhD, MS; Christopher F. Martin, MBA; Yuri Matusov, MD; Alem Mehari, MD; Hector Mendez-Figueroa, MD; Robin Mermelstein, PhD; Torri D. Metz, MD, MS; Richard Morse, BA; Jarrod Mosier, MD; Christian Mouchati, MD; Janet Mullington, PhD; Shawn N. Murphy, MD, PhD; Robert B. Neuman, MD; Janko Z. Nikolich, MD, PhD; Ighovwerha Ofotokun, MD; Elizabeth Ojemakinde, MD, MPH; Anna Palatnik, MD; Kristy Palomares, MD, PhD; Tanyalak Parimon, MD; Samuel Parry, MD; Jan E. Patterson, MD; Thomas F. Patterson, MD; Rachel E. Patzer, PhD, MPH; Michael J. Peluso, MD; Priscilla Pemu, MD, MS; Christian M. Pettker, MD; Beth A. Plunkett, MD, MPH; Kristen Pogreba-Brown, PhD; Athena Poppas, MD; John G. Quigley, MD; Uma Reddy, MD; Rebecca Reece, MD; Harrison Reeder, PhD; W. B. Reeves, MD; Eric M. Reiman, MD; Franz Rischard, DO, MSc; Jonathan Rosand, MD, MS; Dwight J. Rouse, MD; Adam Ruff, BS; George Saade, MD; Grecio J. Sandoval, PhD; Shannon M. Schlater, MS; Fitzgerald Shepherd, MD; Zaki A. Sherif, PhD; Hyagriv Simhan, MD; Nora G. Singer, MD; Daniel W. Skupski, MD; Amber Sowles, RN, BSN; Jeffrey A. Sparks, MD, MMSc; Fatima I. Sukhera, MD; Barbara S. Taylor, MD; Larissa Teunis, MPA; Robert J. Thomas, MD; John M. Thorp, MD, MS; Paul Thuluvath, MD; Amberly Ticotsky, MPH, RN; Alan T. Tita, MD, PhD; Katherine R. Tuttle, MD; Alfredo E. Urdaneta, MD; Daisy Valdivieso, BS; Timothy M. VanWagoner, PhD; Andrew Vasey, MD; Monica Verduzco-Gutierrez, MD; Zachary S. Wallace, MD; Honorine D. Ward, MD; David E. Warren, PhD; Steven J. Weiner, MS; Shelley Welch, MS; Sidney W. Whiteheart, PhD; Zanthia Wiley, MD; Juan P. Wisnivesky, MD, DrPH; Lynn M. Yee, MD; Sokratis Zisis, MD.

Affiliations of RECOVER Consortium Authors: Massachusetts General Hospital, Boston (Alba, Bassett, Bind, Chan, Chibnik, Morse, Murphy, Reeder, Rosand, Wallace); Harvard Medical School, Boston, Massachusetts (Lanca, Mullington, Plunkett); University of Colorado Anschutz Medical Campus, Aurora (Altman, Erlandson, Feuerriegel); Brigham and Women's Hospital, Boston, Massachusetts (Baslet, Bhattacharyya, Sparks); University of Alabama at Birmingham (Erdmann); Case Western Reserve University, Cleveland, Ohio (Zisis); Icahn School of Medicine at Mount Sinai, New York, New York (Charney, Wisnivesky); Stanford University School of Medicine, Stanford, California (Go, Kumar, Urdaneta); Emory University School of Medicine, Atlanta, Georgia (Bedi, Cribbs, Han, Ofotokun); New York University Grossman School of Medicine, New York (Katz); University of Washington, Seattle (Alicic, Franko, Kemp, Lambert, Tuttle); University of California, San Francisco (Anglin, Argueta, Deeks, Grebe, Huang, Peluso, Valdivieso); Howard University, Washington, DC (Ashktorab); Bateman Horne Center, Salt Lake City, Utah (Bateman); Stanford University, Stanford, California (Blomkalns, Bonilla, Geng, Jagannathan); Kaiser Foundation Health Plan of Georgia Inc, Atlanta (Bush, Neuman); University of Kansas Medical Center, Kansas City (Castro); Cedars-Sinai Medical Center, Los Angeles, California (Chen, Matusov); University of Washington School of Medicine, Seattle (Chu); George Washington University, Washington, DC (Clifton); The Ohio State University, Columbus (Costantine); Universidad de Puerto Rico Recinto de Ciencias Medicas, San Juan, Puerto Rico (Davila Nieves); Swedish Medical Center, Seattle, Washington (Duven, Goldman); MaineHealth, Portland (Emery); The University of Arizona, Tucson (Ernst, Lutrick, Nikolich, Pogreba-Brown); Emory University, Atlanta, Georgia (Farah-Abraham, Marconi, Martin, Patzer, Teunis, Wiley); The University of Texas Health Science Center at San Antonio (Farner, J. E. Patterson, T. F. Patterson, Taylor, Verduzco-Gutierrez); University of Illinois Chicago (Fleurimont, Mermelstein); Tulane University Health Sciences Center, New Orleans, Louisiana (Fonseca); Partners HealthCare Systems, Boston, Massachusetts (Gainer); Kaiser Permanente Georgia, Atlanta (Gander); Denver Health, Denver, Colorado (Gardner); MetroHealth Medical Center, Cleveland, Ohio (Gibson, Singer); Pennington Biomedical Research Center, Baton Rouge, Louisiana (Greenway, Kirwan); TriHealth, Cincinnati, Ohio (Habli); University of Illinois Chicago College of Medicine (Hafner); University of Illinois College of Medicine at Peoria (Hanson, S. W. Kelly); Institute for Systems Biology, Seattle, Washington (Heath); UH Cleveland Medical Center, Cleveland, Ohio (Hernandez); University of Utah Schools of the Health Sciences, Salt Lake City (Hess); West Virginia Clinical and Translational Science Institute, Morgantown (Hodder); Christiana Care Health Services Inc, Newark, Delaware (Hoffman); Sanford Health, Sioux Falls, South Dakota (Hoover); Duke University, Durham, North Carolina (Hughes); Cambridge Health Alliance, Cambridge, Massachusetts (John, Ticotsky); Tufts Medical Center, Boston, Massachusetts (Jordan, Ward); MetroHealth Campus of Case Western Reserve University, Cleveland, Ohio (Kaufman); University of California, San Francisco (J. D. Kelly); University of Illinois Chicago (Klein, Quigley); The University of Arizona College of Medicine, Phoenix (Knox, Lee-Iannotti); University of Illinois Hospital and Health Sciences System, Chicago (Krishnan); Howard University College of Medicine, Washington, DC (Laiyemo, Mehari, Sherif); Tulane University, New Orleans, Louisiana (Logarbo); Tulane School of Medicine, New Orleans, Louisiana (Longo); Universidad de Puerto Rico, San Juan, Puerto Rico (Luciano); Beth Israel Deaconess Medical Center, Boston, Massachusetts (Maley, Thomas); Boston University, Boston, Massachusetts (Marathe); University of Mississippi, Oxford (Marshall); McGovern Medical School at The University of Texas Health Science Center at Houston (Mendez-Figueroa); University of Utah Health, Salt Lake City (Metz, Schlater, Sowles); University of Arizona, Tucson (Mosier); Case Western Reserve University School of Medicine, Cleveland, Ohio (Mouchati); Morehouse School of Medicine, Atlanta, Georgia (Ojemakinde, Pemu); Medical College of Wisconsin, Milwaukee (Palatnik); Saint Peter's University Hospital, Brunswick, New Jersey (Palomares); Cedars-Sinai Health System, Los Angeles, California (Parimon); University of Pennsylvania, Philadelphia (Parry); Yale School of Medicine, New Haven, Connecticut (Pettker); NorthShore University HealthSystem, Evanston, Illinois (Plunkett); Warren Alpert Medical School, Brown University, Providence, Rhode Island (Poppas); Columbia University Irving Medical Center, New York, New York (Reddy); West Virginia University School of Medicine, Morgantown (Reece); Department of Medicine, The University of Texas Health Science Center at San Antonio (Reeves); Banner Alzheimer's Institute, Phoenix, Arizona (Reiman); Banner University Medical Center Tucson, Arizona (Rischard); Brown University, Providence, Rhode Island (Rouse); The University of Kansas Medical Center, Kansas City (Ruff); Eastern Virginia Medical School, Norfolk (Saade); Milken Institute of Public Health, The George Washington University, Washington, DC (Sandoval); Boston University School of Medicine, Boston, Massachusetts (Shepherd); University of Pittsburgh, Pittsburgh, Pennsylvania (Simhan); Weill Cornell Medicine, New York, New York (Skupski); The University of Oklahoma, Norman (Sukhera); The University of North Carolina at Chapel Hill (Thorp); Mercy Medical Center, Baltimore, Maryland (Thuluvath); University of Alabama, Birmingham (Tita); University of Oklahoma Health Sciences Center, Oklahoma City (VanWagoner); University of Nebraska Medical Center, Omaha (Vasey, Warren); The George Washington University Biostatistics Center, Rockville, Maryland (Weiner); West Virginia University, Morgantown (Welch); University of Kentucky, Lexington (Whiteheart); Feinberg School of Medicine, Northwestern University, Chicago, Illinois (Yee).

Author Contributions: Drs Thaweethai and Foulkes had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Thaweethai, Jolley, Karlson, Levy, McComsey, Parthasarathy, Singh, Shinnick, Altman, Bassett, Chan, Charney, Chibnik, Clifton, Deeks, Erdmann, Ernst, Fonseca, Gander, Gibson, Heath, Hodder, Hughes, Jagannathan, Jordan, Katz, J. Kelly, Kirwan, Knox, Krishnan, Kumar, Laiyemo, Lee-Iannotti, Maley, Marshall, Mehari, Mendez-Figueroa, Mermelstein, Metz, Mosier, Mullington, Murphy, Nikolich, Ofotokun, J. Patterson, Patzer, Peluso, Pemu, Pettker, Poppas,

Reeves, Reiman, Sherif, Teunis, Thomas, Tita, Verduzco-Gutierrez, Wisnivesky, Yee, Zisis, Horwitz, Foulkes.
Acquisition, analysis, or interpretation of data: Thaweethai, Jolley, Karlson, Levitan, Levy, McComsey, McCorkell, Nadkarni, Parthasarathy, Singh, Walker, Selvaggi, Shinnick, Schulte, Atchley-Challenner, Alba, Alicic, Anglin, Ashktorab, Argueta, Baslet, Bassett, Bateman, Bedi, Bhattacharyya, Bind, Blomkalns, Bonilla, Bush, Castro, Chan, Charney, Chen, Chibnik, Chu, Clifton, Costantine, Cribbs, Davila Nieves, Deeks, Duven, Emery, Erdmann, Erlandson, Farah-Abraham, Farner, Feuerriegel, Fleurimont, Fonseca, Franko, Gainer, Gander, Gardner, Geng, Go, Goldman, Grebe, Greenway, Habli, Hafner, Han, Hanson, Heath, Hernandez, Hess, Hodder, Hoffman, Hoover, Huang, John, Katz, Kaufman, J. Kelly, S. Kelly, Kemp, Kirwan, Klein, Laiyemo, Lambert, Lanca, Lee-Iannotti, Logarbo, Longo, Luciano, Lutrick, Maley, Marathe, Marconi, Martin, Matusov, Metz, Morse, Mosier, Mouchati, Mullington, Murphy, Neuman, Ofotokun, Ojemakinde, Palatnik, Palomares, Parimon, Parry, T. Patterson, Patzer, Peluso, Pemu, Plunkett, Pogreba-Brown, Quigley, Reddy, Reece, Reeder, Rischard, Rosand, Rouse, Ruff, Saade, Sandoval, Schlater, Shepherd, Sherif, Simhan, Singer, Skupski, Sowles, Sparks, Sukhera, Taylor, Teunis, Thomas, Thorp, Thuluvath, Ticotsky, Tita, Tuttle, Urdaneta, Valdivieso, VanWagoner, Vasey, Verduzco-Gutierrez, Wallace, Ward, Warren, Weiner, Welch, Whiteheart, Wiley, Wisnivesky, Yee, Zisis, Horwitz, Foulkes.
Drafting of the manuscript: Thaweethai, Jolley, Karlson, McComsey, McCorkell, Parthasarathy, Singh, Walker, Selvaggi, Bedi, Farner, Fleurimont, Gander, Jordan, J. Kelly, Kirwan, Kumar, Taylor, Teunis, Zisis, Foulkes.
Critical revision of the manuscript for important intellectual content: Thaweethai, Jolley, Karlson, Levitan, Levy, McComsey, McCorkell, Nadkarni, Parthasarathy, Walker, Selvaggi, Shinnick, Schulte, Atchley-Challenner, Alba, Alicic, Altman, Anglin, Ashktorab, Argueta, Baslet, Bassett, Bateman, Bhattacharyya, Bind, Blomkalns, Bonilla, Bush, Castro, Chan, Charney, Chen, Chibnik, Chu, Clifton, Costantine, Cribbs, Davila Nieves, Deeks, Duven, Emery, Erdmann, Erlandson, Ernst, Farah-Abraham, Feuerriegel, Fonseca, Franko, Gainer, Gander, Gardner, Geng, Gibson, Go, Goldman, Grebe, Greenway, Habli, Hafner, Han, Hanson, Heath, Hernandez, Hess, Hodder, Hoffman, Hoover, Huang, Hughes, Jagannathan, John, Jordan, Katz, Kaufman, J. Kelly, S. Kelly, Kemp, Kirwan, Klein, Knox, Krishnan, Laiyemo, Lambert, Lanca, Lee-Iannotti, Logarbo, Longo, Luciano, Lutrick, Maley, Marathe, Marconi, Marshall, Martin, Matusov, Mehari, Mendez-Figueroa, Mermelstein, Metz, Morse, Mosier, Mouchati, Mullington, Murphy, Neuman, Nikolich, Ofotokun, Ojemakinde, Palatnik, Palomares, Parimon, Parry, J. Patterson, T. Patterson, Patzer, Peluso, Pemu, Pettker, Plunkett, Pogreba-Brown, Poppas, Quigley, Reddy, Reece, Reeder, Reeves, Reiman, Rischard, Rosand, Rouse, Ruff, Saade, Sandoval, Schlater, Shepherd, Sherif, Simhan, Singer, Skupski, Sowles, Sparks, Sukhera, Thomas, Thorp, Thuluvath, Ticotsky, Tita, Tuttle, Urdaneta, Valdivieso, VanWagoner, Vasey, Verduzco-Gutierrez, Wallace, Ward, Warren, Weiner, Welch, Whiteheart, Wiley, Wisnivesky, Yee, Zisis, Horwitz, Foulkes.
Statistical analysis: Thaweethai, Selvaggi, Shinnick, Schulte, Bind, Chan, Chibnik, Reeder, Taylor, Foulkes.
Obtained funding: Thaweethai, Karlson, Levy, McComsey, Parthasarathy, Bassett, Charney, Chu, Cribbs, Deeks, Erlandson, Ernst, Heath, Hess, Katz, J. Kelly, Kirwan, Knox, Krishnan, Maley, Martin, Metz, Murphy, Nikolich, Parry, T. Patterson, Patzer, Peluso, Pemu, Quigley, Reeves, Reiman, Saade, Sherif, Simhan, Taylor, Teunis, Tita, Tuttle, Warren, Wisnivesky, Horwitz, Foulkes.
Administrative, technical, or material support: Levy, McComsey, Nadkarni, Singh, Anglin, Ashktorab, Argueta, Bassett, Bedi, Bhattacharyya, Bush, Chan, Charney, Clifton, Costantine, Cribbs, Davila Nieves, Emery, Farah-Abraham, Farner, Feuerriegel, Fleurimont, Franko, Gainer, Gander, Gardner, Geng, Go, Grebe, Habli, Hafner, Han, Hanson, Heath, Hernandez, Hess, Hodder, Hoffman, Huang, Jagannathan, John, Katz, J. Kelly, S. Kelly, Kirwan, Klein, Krishnan, Kumar, Laiyemo, Lambert, Lanca, Lee-Iannotti, Luciano, Lutrick, Maley, Marshall, Martin, Mehari, Mendez-Figueroa, Metz, Morse, Mosier, Mouchati, Mullington, Murphy, Neuman, Ojemakinde, Palomares, Parimon, J. Patterson, Patzer, Peluso, Pettker, Pogreba-Brown, Poppas, Reiman, Rouse, Saade, Schlater, Shepherd, Simhan, Sowles, Sparks, Sukhera, Taylor, Teunis, Thuluvath, Tita, Urdaneta, Valdivieso, Wallace, Ward, Weiner, Yee, Horwitz, Foulkes.
Supervision: Thaweethai, Jolley, Karlson, Levy, McComsey, Parthasarathy, Singh, Walker, Alba, Bassett, Bateman, Bedi, Blomkalns, Bonilla, Castro, Chan, Charney, Chen, Chu, Costantine, Emery, Erlandson, Farah-Abraham, Gander, Goldman, Greenway, Hafner, Han, Heath, Hernandez, Hodder, Hoffman, Huang, Jagannathan, Jordan, Katz, J. Kelly, Kirwan, Knox, Krishnan, Lee-Iannotti, Longo, Martin, Mendez-Figueroa, Metz, Mullington, Murphy, Neuman, Nikolich, Ofotokun, Patzer, Peluso, Plunkett, Poppas, Reiman, Rosand, Saade, Shepherd, Simhan, Skupski, Sukhera, Thorp, Thuluvath, Tita, Ward, Warren, Whiteheart, Yee, Horwitz, Foulkes.
Other - gave feedback based on study design: Singer.
Other - cohort recruitment and study supervision: Alba.
Other - was one of the original Boston Hub investigators involved in application for funding: Mullington.
Other - operationalization: Farah-Abraham.
Other - literature review: Atchley-Challenner.
Other - site PI team management: Marconi.
Other - data procurement: Kumar.

Conflict of Interest Disclosures: Dr Thaweethai reported receiving grants from the National Institutes of Health (NIH)/National Heart, Lung, and Blood Institute (NHLBI) and serving as co-investigator of the RECOVER Data Resource Core. Dr Jolley reported receiving grants from the NIH/NHLBI. Dr Karlson reported receiving grants from the NIH. Dr Levitan reported receiving grants from the NIH and Amgen. Dr Levy reported receiving grants from the NHLBI and personal fees or research support from Entrinsic Bioscience, Nocion Therapeutics, Gossamer Bio, AstraZeneca, Novartis, Sanofi, Amgen, Genentech, Pieris Pharmaceuticals, and SRA. Dr McComsey reported serving as a consultant for Gilead, Merck, Janssen, and GlaxoSmithKline and receiving research support from Pfizer, Genentech, and Roche. Ms McCorkell reported receiving personal fees from NIH RECOVER and research funding as a subcontractor from the NIH and the Patient-Centered Outcomes Research Institute (PCORI). Dr Parthasarathy reported receiving grants from the NIH, NHLBI, Sergey-Brin Foundation, and Regeneron and personal fees from Jazz Pharmaceuticals. Dr Singh reported receiving grants from Pfizer and serving as an advisor to Regeneron and Gilead. Dr Walker reported receiving grants from the NHLBI. Mr Shinnick reported receiving grants from the NIH/NHLBI. Dr Horwitz reported receiving grants from the NIH and serving on an ad hoc committee for the National Academy of Medicine. Dr Foulkes reported receiving grants from the NIH/NHLBI and personal fees from the Round Table Group and serving as principal investigator of the RECOVER Data Resource Core. Dr Alba reported receiving personal fees from Minerva Biotechnologies. Dr Alicic reported receiving grants from the NIH, the Centers for Disease Control and Prevention, and Travere Therapeutics and personal fees from Boehringer Ingelheim. Dr Ashktorab reported receiving grants from the NIH. Dr Bhattacharyya reported receiving grants from NIH RECOVER, Alexion Pharmaceuticals, UCB, and Massachusetts Consortium on Pathogen Readiness and personal fees from American Academy of Neurology and UpToDate. Dr Castro reported receiving grants from the NIH, American Lung Association, PCORI, AstraZeneca, Gala Therapeutics, Genentech, GlaxoSmithKline, Novartis, Pulmatrix, Sanofi, Shionogi, and Theravance Biopharma; personal fees from Allakos, Amgen, OM Pharma, Pfizer, Pioneering Medicines, AstraZeneca, GlaxoSmithKline, Genentech, Teva Pharmaceutical Industries, Sanofi, Merck, Novartis, Arrowhead Pharmaceuticals, and Regeneron; stock options from Aer Therapeutics; and royalties from Elsevier. Dr Chu reported receiving grants from the NIH and personal fees from Merck, Vir Biotechnology, AbbVie, Ellume, Pfizer, and the Bill and Melinda Gates Foundation; serving on advisory boards for Merck and AbbVie; conducting CME teaching with Medscape, Vindico, and Clinical Care Options; and receiving research funding from Gates Ventures and support and reagents from Ellume and Cepheid. Dr Clifton reported receiving grants from the NIH. Dr Costantine reported receiving grants from the NIH, Baxter International, and Siemens Healthcare and personal fees from Quidel, Progenity, and Siemens Healthcare. Dr Davila Nieves reported receiving grants from the NIH. Dr Deeks reported receiving grants from the NIH. Dr Emery reported receiving grants from the NIH. Dr Erdmann reported receiving grants from the NIH and personal fees from Perspectum and having a patent for Plantform. Dr Erlandson reported receiving grant funding and advisory fee payments from Gilead Sciences (paid to the University of Colorado) and advisory fee payments from Merck and ViiV (paid to the University of Colorado) in the last 3 years. Ms Gainer reported receiving grants from NIH RECOVER. Dr Geng reported receiving grants from Pfizer and personal fees from UnitedHealthcare. Dr Gibson reported receiving grants from the NIH. Dr Go reported receiving grants from the NIH. Dr Goldman reported receiving grants from Gilead Sciences and Merck; personal fees from Gilead Sciences and Eli Lilly; contracted research from Gilead Sciences, Eli Lilly, and Regeneron Pharmaceuticals; and nonfinancial support from Adaptive Biotechnologies, Labcorp, and Monogram Biosciences. Dr Greenway reported receiving grants from the NIH. Dr Heath reported receiving grants

from Merck and Gilead Sciences, personal fees from Regeneron Pharmaceuticals, and being a board member for IsoPlexis. Dr Hess reported receiving grants from the NIH and being a member of a data and safety monitoring board for Astellas Pharmaceuticals. Dr Hodder reported receiving grants from the NHLBI. Dr Hoffman reported receiving grants from the NHLBI. Dr Hughes reported receiving honorarium from UpToDate and personal fees from AMAG Pharmaceuticals. Ms John reported receiving grants from the NIH. Dr Katz reported receiving grants from the NIH. Dr Klein reported receiving grants from NIH RECOVER and consulting fees from Gilead Sciences. Dr Knox reported receiving grants from the NIH. Dr Krishnan reported receiving grants from the NIH/NHLBI; personal fees from GlaxoSmithKline, AstraZeneca, CereVu Medical, BData, Propeller, ResMed, American Board of Internal Medicine, American Academy of Allergy, Asthma, and Immunology; and research funding from the NIH, PCORI, American Lung Association, COPD Foundation, and the Sergey Brin Family Foundation. Dr Laiyemo reported receiving grants from the NIH. Dr Lambert reported receiving grants from the NIH and the PCORI and serving as a site principal investigator for trials funded by Vertex and Aceragan. Dr Luciano reported receiving grants from the NIH and speaker fees from Sanofi-Genzyme. Dr Marathe reported receiving grants from NIH RECOVER and funding from Boston Medical Center. Dr Marconi reported receiving grants from the Centers for Disease Control and Prevention, Veteran Affairs, and the NIH; grants, personal fees, nonfinancial support, and other from Eli Lilly and Gilead; grants and personal fees from ViiV; and nonfinancial support from Bayer. Dr Marshall reported receiving grants from the NIH. Mr Martin reported receiving grants from the NIH. Dr Metz reported receiving grants from the NIH and Pfizer and personal fees from Pfizer. Mr Morse reported receiving grants from the NIH/NHLBI. Dr Mullington reported receiving grants from the NIH and Open Medicine Foundation and speaker and book chapter contribution for Idorsia Pharmaceuticals. Dr Murphy reported receiving grants from the NIH/NHLBI. Dr Nikolich reported receiving grants from the NIH/NHLBI. Dr Ojemakinde reported receiving grants from the NIH. Dr Parry reported receiving grants from NIH RECOVER paid to Penn and funding from the Sergey Brin Family Foundation. Dr Peluso reported receiving personal fees from Gilead Sciences and AstraZeneca. Dr Pogreba-Brown reported receiving grants from NIH RECOVER. Dr Poppas reported owning stock in GE and serving as an UpToDate contributor and co-editor of Hurst and Fuster's The Heart. Dr Quigley reported receiving grants from the NIH, Pfizer, AbbVie, and Teva Pharmaceutical Industries and personal fees from Alexion Pharmaceuticals, Servier Laboratories, Agios Pharmaceuticals, and Rigel Pharmaceuticals. Dr Reiman reported receiving grants from the NIH. Dr Rischard reported receiving grants from the NIH/NHLBI, Bayer, Janssen, Merck, Aerovate Therapeutics, and Respira Therapeutics and consulting relationships with Acceleron Pharma and United Therapeutics. Dr Rosand reported receiving grants from the NIH and the American Heart Association and personal fees from the National Football League and Takeda Pharmaceutical Co. Dr Shepherd reported receiving grants from NIH RECOVER. Dr Simhan reported receiving grants from the NIH. Dr Singer reported receiving grants from Case Western Reserve University and MetroHealth. Dr Sparks reported receiving grants from the National Institute of Arthritis and Musculoskeletal and Skin Diseases, the R. Bruce and Joan M. Mickey Research Scholar Fund, the Llura Gund Award for Rheumatoid Arthritis Research and Care, Brigham and Women's Hospital, and Bristol Myers Squibb and personal fees from Bristol Myers Squibb, AbbVie, Amgen, Boehringer Ingelheim, Gilead Sciences, Inova Diagnostics, Janssen, Optum, and Pfizer. Dr Thomas reported receiving personal fees from Guidepoint Global and GLG Councils and having a patent for ECG-spectrogram with royalties paid from MyCardio LLC through Beth Israel Deaconess Medical Center. Dr Thuluvath reported receiving grants from Mercy Medical Center. Dr Tita reported receiving grants from Pfizer. Dr VanWagoner reported receiving grants from the NIH. Dr Vasey reported receiving grants from NIH RECOVER. Dr Wallace reported receiving grants from Bristol Myers Squibb, Sanofi, and Horizon Therapeutics and personal fees from Sanofi, Novartis, Shionogi, Visterra, Horizon Therapeutics, PPD Inc, Zenas BioPharma, and Medpace. Dr Warren reported receiving grants from the NIH. Mr Weiner reported receiving grants from the NIH. Ms Welch reported receiving grants from the NIH. Dr Wisnivesky reported receiving personal fees from PPD Inc, Banook, Sanofi, Atea Pharmaceuticals, and Prospero and grants from Regeneron Pharmaceuticals, Axella, Sanofi, and Arnold. Dr Yee reported receiving grants from the NIH. No other disclosures were reported.

Funding/Support: This research was funded by the NIH (OTA OT2HL161841, OT2HL161847, and OT2HL156812) as part of the Researching COVID to Enhance Recovery (RECOVER) research program. Additional support for Drs Foulkes, Thaweethai, and Schulte was provided by R01 HL162373.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Group Information: The RECOVER Consortium nonauthor collaborators are listed in Supplement 4.

Disclaimer: This content is solely the responsibility of the authors and does not necessarily represent the official views of the RECOVER Initiative, the NIH, or other funders.

Data Sharing Statement: See Supplement 5.

Additional Information: This study is part of the RECOVER initiative, which seeks to understand, treat, and prevent the postacute sequelae of SARS-CoV-2 infection. For more information on RECOVER, visit https://recovercovid.org/. We thank all the participants enrolled in the RECOVER Initiative, the National Community Engagement Group (NCEG), and all patient, caregiver and community representatives.

REFERENCES

1. World Health Organization. WHO coronavirus (COVID-19) Dashboard. Accessed February 28, 2023. https://covid19.who.int/

2. Danesh V, Arroliga AC, Bourgeois JA, et al. Symptom clusters seen in adult COVID-19 recovery clinic care seekers. *J Gen Intern Med*. 2023;38(2):442-449. doi:10.1007/s11606-022-07908-4

3. Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine*. 2021;38:101019. doi:10.1016/j.eclinm.2021.101019

4. FAIR Health. A detailed study of patients with long-haul COVID: an analysis of private healthcare claims. June 15, 2021. Accessed February 28, 2023. https://collections.nlm.nih.gov/catalog/nlm:nlmuid-9918334383006676-pdf

5. Huang C, Huang L, Wang Y, et al. 6-Month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet*. 2021;397(10270):220-232. doi:10.1016/S0140-6736(20)32656-8

6. Reese JT, Blau H, Casiraghi E, et al; N3C Consortium; RECOVER Consortium. Generalisable long COVID subtypes: findings from the NIH N3C and RECOVER programmes. *EBioMedicine*. 2023;87:104413. doi:10.1016/j.ebiom.2022.104413

7. Cutler DM. The costs of long COVID. *JAMA Health Forum*. 2022;3(5):e221809. doi:10.1001/jamahealthforum.2022.1809

8. Perlis RH, Trujillo KL, Safarpour A, et al. Research letter: association between long COVID symptoms and employment status. *medRxiv*. November 18, 2022. doi:10.1101/2022.11.17.22282452

9. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol*. 2023;21(3):133-146. doi:10.1038/s41579-022-00846-2

10. Merad M, Blish CA, Sallusto F, Iwasaki A. The immunology and immunopathology of COVID-19. *Science*. 2022;375(6585):1122-1127. doi:10.1126/science.abm8108

11. World Health Organization. Public health surveillance for COVID-19: interim guidance. July 22, 2022. Accessed February 28, 2023. https://www.who.int/publications/i/item/WHO-2019-nCoV-SurveillanceGuidance-2022.2

12. RECOVER. Research Protocols. Accessed February 28, 2023. https://recovercovid.org/protocols

13. Hastie T, Tibshirani R, Friedman JH, Friedman JH. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Vol 2. Springer; 2009. doi:10.1007/978-0-387-84858-7

14. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach Learn*. 2003;52:91-118. doi:10.1023/A:1023949509487

15. Murtagh F, Contreras P. Algorithms for hierarchical clustering: an overview. *Wiley Interdiscip Rev Data Min Knowl Discov*. 2012;2(1):86-97. doi:10.1002/widm.53

16. Fernández-de-Las-Peñas C, Notarte KI, Peligro PJ, et al. Long-COVID symptoms in individuals infected with different SARS-CoV-2 variants of concern: a systematic review of the literature. *Viruses*. 2022;14(12):2629. doi:10.3390/v14122629

17. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature*. 2021;594(7862):259-264. doi:10.1038/s41586-021-03553-9

18. Estiri H, Strasser ZH, Brat GA, Semenov YR, Patel CJ, Murphy SN; Consortium for Characterization of COVID-19 by EHR (4CE). Evolving phenotypes of non-hospitalized patients that indicate long COVID. *BMC Med*. 2021;19(1):249. doi:10.1186/s12916-021-02115-0

19. Horberg MA, Watson E, Bhatia M, et al. Post-acute sequelae of SARS-CoV-2 with clinical condition definitions and comparison in a matched cohort. *Nat Commun*. 2022;13(1):5822. doi:10.1038/s41467-022-33573-6

20. Mizrahi B, Sudry T, Flaks-Manov N, et al. Long COVID outcomes at one year after mild SARS-CoV-2 infection: nationwide cohort study. *BMJ*. 2023;380:e072529. doi:10.1136/bmj-2022-072529

21. Roessler M, Tesch F, Batram M, et al. Post-COVID-19-associated morbidity in children, adolescents, and adults: a matched cohort study including more than 157,000 individuals with COVID-19 in Germany. *PLoS Med*. 2022;19(11):e1004122. doi:10.1371/journal.pmed.1004122

22. Wulf Hanson S, Abbafati C, Aerts JG, et al; Global Burden of Disease Long COVID Collaborators. Estimated global proportions of individuals with persistent fatigue, cognitive, and respiratory symptom clusters following symptomatic COVID-19 in 2020 and 2021. *JAMA*. 2022;328(16):1604-1615. doi:10.1001/jama.2022.18931

23. Huang Y, Pinto MD, Borelli JL, et al. COVID symptoms, symptom clusters, and predictors for becoming a long-hauler looking for clarity in the haze of the pandemic. *Clin Nurs Res*. 2022;31(8):1390-1398. doi:10.1177/10547738221125632

24. Brannock MD, Chew RF, Preiss AJ, et al. Long COVID risk and pre-COVID vaccination: an EHR-based cohort study from the RECOVER program. *medRxiv*. October 7, 2022. doi:10.1101/2022.10.06.22280795

25. Kuodi P, Gorelik Y, Zayyad H, et al. Association between BNT162b2 vaccination and reported incidence of post-COVID-19 symptoms: cross-sectional study 2020-21, Israel. *NPJ Vaccines*. 2022;7(1):101. doi:10.1038/s41541-022-00526-5

26. Haffke M, Freitag H, Rudolf G, et al. Endothelial dysfunction and altered endothelial biomarkers in patients with post-COVID-19 syndrome and chronic fatigue syndrome (ME/CFS). *J Transl Med*. 2022;20(1):138. doi:10.1186/s12967-022-03346-2

27. Kedor C, Freitag H, Meyer-Arndt L, et al. A prospective observational study of post-COVID-19 chronic fatigue syndrome following the first pandemic wave in Germany and biomarkers associated with symptom severity. *Nat Commun*. 2022;13(1):5104. doi:10.1038/s41467-022-32507-6

28. Larsen NW, Stiles LE, Shaik R, et al. Characterization of autonomic symptom burden in long COVID: a global survey of 2,314 adults. *Front Neurol*. 2022;13:1012668. doi:10.3389/fneur.2022.1012668

29. Mancini DM, Brunjes DL, Lala A, Trivieri MG, Contreras JP, Natelson BH. Use of cardiopulmonary stress testing for patients with unexplained dyspnea post-coronavirus disease. *JACC Heart Fail*. 2021;9(12):927-937. doi:10.1016/j.jchf.2021.10.002

30. Twomey R, DeMars J, Franklin K, Culos-Reed SN, Weatherald J, Wrightson JG. Chronic fatigue and postexertional malaise in people living with long COVID: an observational study. *Phys Ther*. 2022;102(4):pzac005. doi:10.1093/ptj/pzac005