

UCSF

UC San Francisco Previously Published Works

Title

Comparability of control and comparison groups in studies assessing long COVID

Permalink

<https://escholarship.org/uc/item/8077475t>

Authors

Haslam, Alyson

Prasad, Vinay

Publication Date

2023

DOI

10.1016/j.amjmed.2023.01.005

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Comparability of Control and Comparison Groups in Studies Assessing Long COVID

Alyson Haslam, PhD, Vinay Prasad, MD, MPH

University of California San Francisco, San Francisco.

ABSTRACT

BACKGROUND: Awareness of long coronavirus disease (COVID) began primarily through media and social media sources, which eventually led to the development of various definitions based on methodologies of varying quality. We sought to characterize comparison groups in long COVID studies and evaluate comparability of the different groups.

METHODS: We searched Embase, Web of Science, and PubMed for original research articles published in high-impact journals. We included studies on human patients with long COVID outcomes, and we abstracted study-related characteristics, as well as long COVID characteristics.

RESULTS: Of the 83 studies, 3 were randomized controlled trials testing interventions for long COVID, and 80 (96.4%) were observational studies. Among the 80 observational studies, 76 (95%) were trying to understand the incidence, prevalence, and risk factors for long COVID, 2 (2.5%) examined prevention strategies, and 2 (2.5%) examined treatment strategies. Among those 80 studies, 45 (56.2%) utilized a control or comparison group and 35 (43.8%) did not. Compared with 95% of observational studies that documented symptoms or assessed risk factors, all randomized studies assessed treatment strategies. We found 48.8% of observational studies did any adjustment for covariates, including demographics or health status. Of those that did adjust for covariates, 15 (38.5%) adjusted for 4 or fewer variables. We found that 26.5% of all studies and 45.8% of studies with a control/comparator group matched participants on at least 1 variable.

CONCLUSION: Long COVID studies in high-impact journals primarily examine symptoms and risk factors of long COVID; often lack an adequate comparison group and often do not control for potential confounders. Our results suggest that standardized definitions for long COVID, which are often based on data from uncontrolled and potentially biased studies, should be reviewed to ensure that they are based on objective data.

© 2023 Elsevier Inc. All rights reserved. • *The American Journal of Medicine* (2023) 000:1–5

KEYWORDS: control arm; COVID-19; long COVID; long-haulers; study design

INTRODUCTION

Since the emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), there have been reports of

Funding: None.

Conflicts of Interest: AH reports none. VP reports (Research funding) Arnold Ventures; (Royalties) Johns Hopkins Press, Medscape, and MedPage (Honoraria) Grand Rounds/lectures from universities, medical centers, non-profits, and professional societies. (Consulting) UnitedHealthcare and OptumRX. (Other) Plenary Session podcast has Patreon backers, YouTube, and Substack.

Authorship: Both authors had access to the data and a role in writing this manuscript.

Requests for reprints should be addressed to Alyson Haslam, PhD, Department of Epidemiology and Biostatistics, UCSF Mission Bay Campus | Mission Hall: Global Health & Clinical Sciences Building | 550 16th St, 2nd Fl, San Francisco, CA, 94158.

E-mail address: alyson.haslam@ucsf.edu

people with persistent or long-term sequelae from infection, including fatigue, shortness of breath, and cognitive impairment. These reports were initially shared on social media, and writers in major media sources began to dub the condition “long coronavirus disease (COVID)” or describe those suffering as “long-haulers.”^{1,2} Because the origins of long COVID came through social media and not the biomedical literature, some have suggested that this is an illness that doctors only became aware of through empowered patients.¹

Numerous investigations have probed the symptoms and natural history of those with persisting symptoms post-COVID, culminating in the development of several definitions of long COVID.^{3,4} However, these definitions were formulated by pooling study designs and methodologies of

varying quality. This may lead to misclassification and bias in symptoms and prevalence estimates.

The objective of our study was to characterize control and comparison groups and evaluate comparability of the different groups in studies evaluating long COVID or postacute COVID syndrome. Because we were interested in studies with the greatest impact and most rigorous methodology, we chose to focus on studies published in high-impact journals.

MATERIALS AND METHODS

Search Strategy

We searched Web of Science and Embase for the term “long COVID” and restricted to “articles” in English. We conducted a separate search on PubMed, using the terms “long COVID” or “post acute COVID syndrome,” and limited the search to clinical and observational studies. Our search date was October 13, 2022. We included any original research study design that included human patients with long COVID or had a persistent or long-term COVID-related outcome. We excluded animal studies, attitudes about COVID, case reports/case series with <10 people (we included case series of more than 10), basic science/cell, device/algorithm, modeling/simulation, reviews, research letters, studies not examining the condition of long COVID or a health outcome, protocols, qualitative reporting, social media, and test validation. We then excluded studies that were published in journals with an impact factor of less than 10, as per the most recent impact factor published on the journal’s website.

Data Abstraction

For each included study, we abstracted data on the study design, categorized as cross sectional, including surveys, retrospective observational, case control, prospective observational, including pre-/post studies, and randomized controlled trials; study population, categorized as outpatient, inpatient, both inpatient and outpatient, community/general public, health care workers, national database, social media participants, or not indicated; general age of study participants, categorized as adult, child, both, or not indicated; number of people in the study; number in each arm/group (if multiple); whether there was a control arm or comparison group; whether vaccination was taken into account; the dates of enrollment or dates of data inclusion; definition of long COVID (symptoms, duration of symptoms, confirmed/unconfirmed COVID diagnosis); prevalence of long COVID; country of study population; information on intervention/exposure and control/unexposed groups; data on matching and/or adjustment for covariates; whether there were sensitivity analyses to test for negative or positive

control outcomes or exposures; funding information; and conflict of interest disclosures.

To examine the citations and media attention of the articles, we obtained the altmetric score for each article, using the Altmetric it! extension.

CLINICAL SIGNIFICANCE

- In high-impact journals, 48 (57.8%) studies used a comparison group and 35 (42.2%) did not.
- About 50% of studies did any adjustment for confounders, and of those that did adjust for covariates, 42.8% adjusted for four or fewer variables.
- Our results have implications for standardized definitions, which are often based on data from uncontrolled and potentially biased studies.

Statistical Analysis

We calculated descriptive statistics for the overall analytic sample and for the sample, stratified by whether the study used a control or comparator group. We calculated categorical differences with χ^2 tests and continuous variables with Wilcoxon-rank sum tests. We conducted all analyses using R software, version 3.6.2, and Microsoft Excel. In accordance with 45 CFR §46.102(f), this study was not submitted for institutional review

board approval because it involved publicly available data and did not involve individual patient data.

RESULTS

Our search identified 927 results in Embase, 882 in Web of Science, and 61 on PubMed (Figure 1). After removing duplicate studies ($n = 932$), studies not meeting inclusion criteria ($n = 452$), and studies published in journals with impact factors of less than 10 ($n = 407$), 83 articles were included in the analytic sample. Study characteristics are detailed in Supplementary Table 1 (available online), stratified by control or comparison group status.

The median journal impact factor for included studies was 18 (interquartile range [IQR]: 12, 21). The median number of participants per study was 599 (IQR: 150, 7584), and the median age was 49 (IQR: 45, 56; $n=71$ studies). The year that studies most commonly completed enrollment was in 2020 (65.1%; $n = 54$). Most studies were conducted in a European country (68.7%; $n = 57$). Most studies were funded by a nonindustry source (69.9%; $n = 58$), and most authors reported either payments from industry (47.0%; $n = 39$) or no conflict of interest (45.8%; $n = 38$). Frequencies of studies published by year, study design type, and including a comparator group are shown in Figure 2.

Of the 83 studies, 48 (57.8%) used a control or comparison group, and 35 (42.2%) did not. Compared to studies without a control/comparator group, studies with a control/comparator group were more likely to use a retrospective observational study design (25.0% vs 5.7%) and less likely to use a prospective cohort (45.8% vs 68.6%; $P = .007$ for global differences). Compared to studies without a control/comparator group, studies with a control/comparator group were more likely to use a population that included outpatients (47.9% vs 28.6%) and less likely to use a population from the general community (8.3% vs 22.9%) or inpatients (12.5% vs 31.4%; $P = .04$ for global differences).

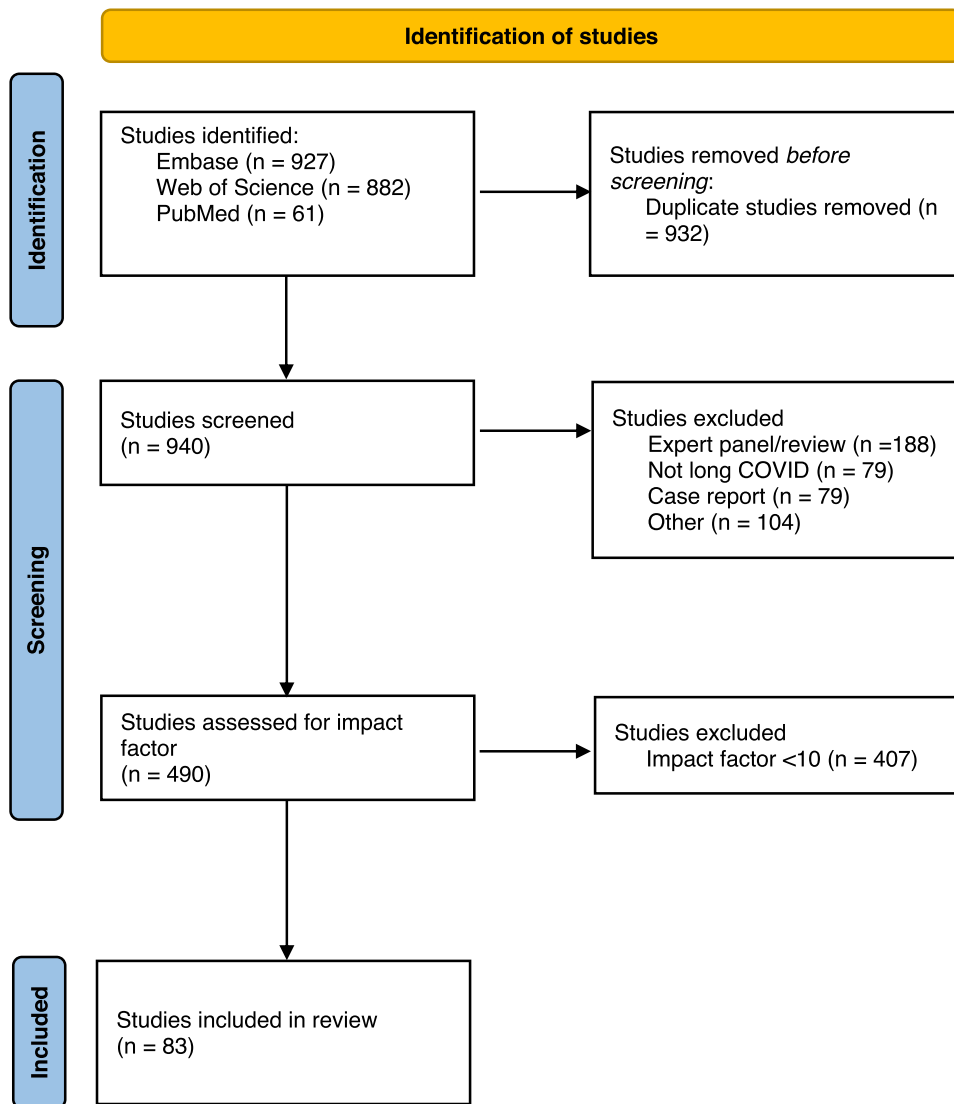


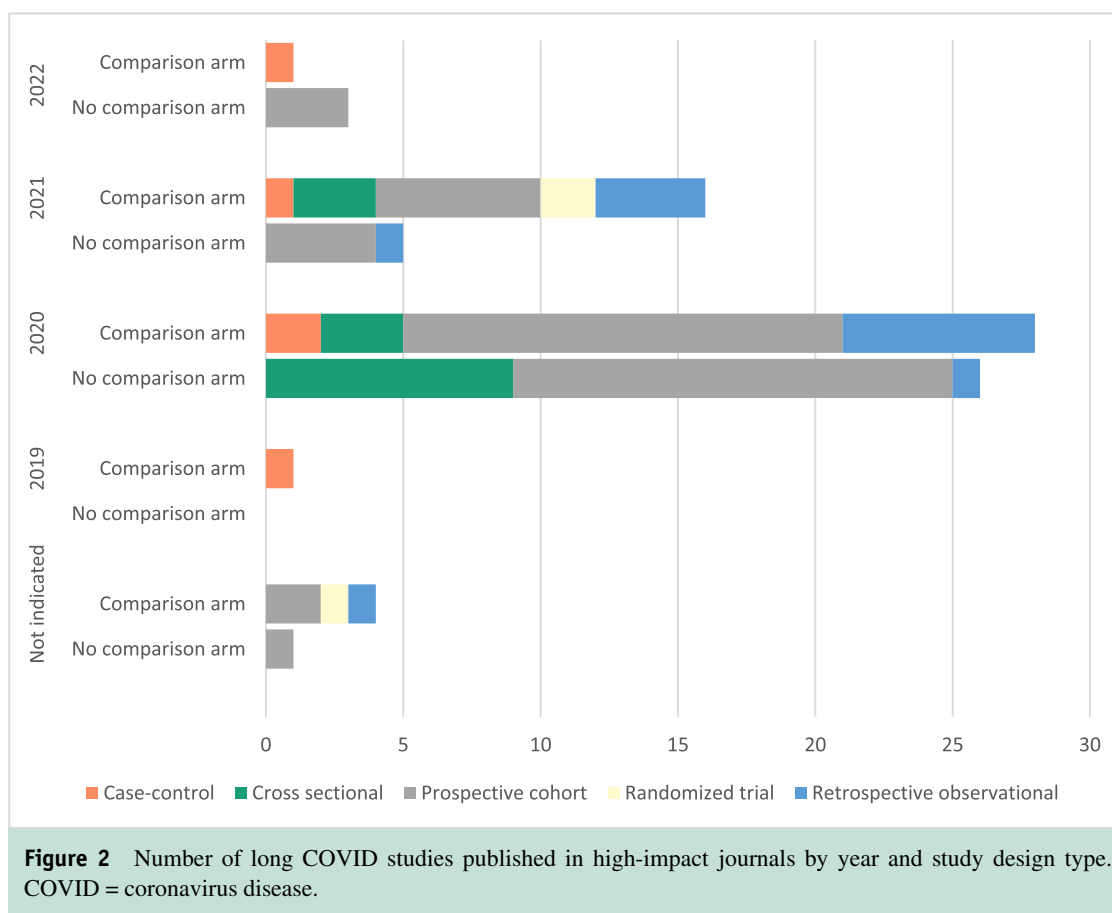
Figure 1 Flowchart for selecting studies that publish findings on long COVID in high-impact medical journals. COVID = coronavirus disease.

Supplementary Table 2 (available online), lists COVID definition characteristics and control/ comparator group characteristics, by comparator group status and for the 3 randomized studies. Compared to studies without a control/comparator group, studies with a control/comparator group were less likely to have a nonspecific list of long COVID symptoms (41.7% vs 68.6%; $P = .03$) and a lower median number of possible symptoms (18 vs 21; $P = .05$). The most common COVID symptom duration for studies without a control/comparator group was not indicated (34.3%), whereas it was 3-4 weeks (35.4%) for studies with a control/comparator group.

Most studies did not account for vaccinations in their analysis, but most studies were conducted before COVID vaccines were available (63.9%; $n = 53$), while 9.6% ($n = 8$) were completed postvaccine and did not account for vaccinations in their analysis. About 50% ($n = 42$) of studies did any adjustment for covariates (including

demographics or health status), and of those that did adjust for covariates, 42.8% ($n = 18/42$) adjusted for 4 or fewer variables. Of all studies, 26.5% and 45.8% of studies ($n = 22$) with a control/comparator group matched participants on at least 1 variable. Age ($n = 13$) and sex ($n = 12$) were the most commonly reported variables used for matching.

Of studies that did have a control/comparator group ($n = 48$), 8.3% ($n = 4$) used a negative control outcome sensitivity analysis; 6.2% ($n = 3$) used a positive control outcome sensitivity analysis; and 2.1% ($n = 1$) used negative exposure outcome sensitivity analysis. Of studies with a control/comparator group, 47.9% ($n = 23$) reported a comparison of baseline demographic and health status variables between groups, while 20.8% ($n = 10$) compared baseline demographics with limited health outcomes, 12.5% ($n = 6$) compared baseline demographics, and 18.8% ($n = 9$) presented no baseline comparisons between groups.



The median prevalence of long COVID was 48 (IQR: 36, 66). When stratifying by the inclusion of a comparison group, the prevalence of long COVID among those infected with COVID was 51 (IQR: 42, 68) for studies without a comparison group and 46 (IQR: 14, 65) for studies without a comparison group. The prevalence of long COVID symptoms in comparator group participants was 34 (IQR: 7, 38).

The 3 randomized studies in the analytic sample examined breathing interventions to treat breathlessness in people with COVID ($n = 2$) and a monoclonal antibody to treat people with long COVID ($n = 1$), whereas nonrandomized studies were primarily focused on documenting symptoms ($n = 52$ out of 80; 65.0%) or risk factors ($n = 24$ out of 80; 30.0%). Two nonrandomized studies examined prevention of long COVID, and 2 studies examined intervention exposures aimed at preventing long COVID symptoms.

DISCUSSION

We found that in studies reporting on long COVID in high-impact journals, about 40% of studies assessing long COVID did not include a control or comparison group. We also found that about half of these studies do not adjust for variables that may also be associated with long COVID symptoms, such as comorbidities and age, and when studies do adjust for such covariates, most adjust for only a few

variables. Further, only a small handful of these studies (<5%) perform sensitivity analyses. Finally, we found, as others have, that the diagnosis, list of symptoms, and duration of symptoms of long COVID vary.⁵

In assessing the prevalence and natural history of long COVID, including a control or comparison group is an important step in ensuring that symptoms of long COVID are not a result of some other personal, social, or environmental characteristic. These characteristics could include aging, health status, or policy implementation, which are unrelated to the specific COVID infection but may predispose an individual to being exposed to or acquiring severe acute respiratory syndrome coronavirus 2. The COVID pandemic has affected people's lives both through direct viral impact as well as disruption to social systems and routines. This combination has led to changes in behavior, such as physical activity, sleep, and in-person and virtual interactions,⁶ and biologic changes, even among those not infected with COVID.⁷

The definition of long COVID has evolved throughout the pandemic, prompting several large organizations to develop differing definitions with broad inclusion criteria.^{3,4} Because many of the studies in our analytic sample were conducted prior to the publication of these definitions, at least some of these studies were used in the development of these well-recognized definitions. When definitions are developed, especially ones that will be used on a large-scale level, they

should be based upon unbiased and objective data, and yet, we find that if studies comparable to the studies in our analytic cohort are used in the development of long COVID definitions, there is the potential for bias and misclassification in determining long COVID.

In our study, we included all study designs, and the best study design will depend on the research question being studied. For questions assessing efficacy of an intervention, such as vaccines and treatment strategies, a randomized study is essential to ensure comparability between study arms. Of the 83 studies in our analytic sample, 3 were randomized studies, testing interventions to treat long COVID or its symptoms.

Alternatively, studies that assess claims about COVID natural history and severity cannot be randomized. Therefore, it is important to include comparison groups, particularly those that are analogous to the exposure group. Most of the nonrandomized studies in our analytic sample (95%) examined symptoms or risk factors for long COVID. However, a limitation exists because these studies are not randomized and, therefore, require further steps to minimize confounding between exposure groups. For example, a recent prospective observational study examined physical, mental, and social well-being outcomes between patients with COVID-19 and other upper respiratory infections.⁸ This particular study included a control group, consisting of those with upper respiratory infections, which was more comparable than healthy people to those with COVID-19 infections. By doing this, researchers were better able to determine whether lingering symptoms were attributed to COVID-19 specifically or whether those symptoms can be attributed to any respiratory infections.

In contrast, a recently published economic analysis did not provide a comparison of costs between COVID and other respiratory tract infections. Their analysis estimated the societal cost of long COVID to be valued at \$3.7 trillion, yet it did not provide a comparison of costs related to other respiratory tract infections.⁹

Strengths and Limitations

A strength of this study is that this is the first broad analysis of literature that seeks to examine comparison measures in long COVID studies. Second, we focused on high-impact publications, which have played an inordinate role in steering the public conversation. The median altmetric score of articles we examined was 173. Moreover, these journals are well known for strong peer-review process with accomplished reviewers and editors.

We recognize that there are several limitations to our analysis. First, the studies we included in this analysis cannot be generalized to others as we used publication in high-impact journals as an inclusion criterion. Because we used studies from high-impact journals, it is likely that our analytic sample includes studies with even higher study quality

and less bias than studies published in lower impact factor journals. Second, we relied on information published in the studies, so some variables in the studies may have been misclassified because of lack of reporting due to word limits or oversight. Third, because of a limited number of studies, we were not able to examine the interaction of different variables, such as year, study type, and study population. Fourth, because there were only 3 randomized studies, we were unable to do an in-depth comparison of differences in study methodologies and outcomes between randomized and nonrandomized studies.

CONCLUSION

In conclusion, we found that more than 40% of long COVID studies published in high-impact journals lack a comparator group and that many studies do not adjust for (49%) or match on (74%) confounding variables. Standardized definitions for long COVID, which are often based on uncontrolled data and potentially biased studies, should be reviewed to ensure that they are based on objective data.

References

- Callard F, Peregó E. How and why patients made long COVID. *Soc Sci Med* 2021;268:113426. <https://doi.org/10.1016/j.socscimed.2020.113426>.
- Yong E. Long-Haulers Are Redefining COVID-19. *The Atlantic*. Available at: <https://www.theatlantic.com/health/archive/2020/08/long-haulers-covid-19-recognition-support-groups-symptoms/615382/>. Accessed October 19, 2022.
- National Institute for Health and Care Excellence (NICE). *COVID-19 Rapid Guideline: Managing the Long-Term Effects of COVID-19*. London, UK: NICE; 2020.
- World Health Organization (WHO). A Clinical Case Definition of Post COVID-19 Condition by a Delphi Consensus. Available at: https://www.who.int/publications/i/item/WHO-2019-nCoV-Post_COVID-19_condition-Clinical_case_definition-2021.1. Accessed September 5, 2022.
- Ceban F, Leber A, Jawad MY, et al. Registered clinical trials investigating treatment of long COVID: a scoping review and recommendations for research. *Infect Dis (Lond)* 2022;54(7):467–77. <https://doi.org/10.1080/23744235.2022.2043560>.
- Giuntella O, Hyde K, Saccardo S, Sadoff S. Lifestyle and mental health disruptions during COVID-19. *Proc Natl Acad Sci USA* 2021;118(9):e2016632118. <https://doi.org/10.1073/pnas.2016632118>.
- Brusaferri L, Alshelhi Z, Martins D, et al. The pandemic brain: neuroinflammation in non-infected individuals during the COVID-19 pandemic. *Brain Behav Immun* 2022;102:89–97. <https://doi.org/10.1016/j.bbi.2022.02.018>.
- Wisk LE, Gottlieb MA, Spatz ES, et al. Association of initial SARS-CoV-2 test positivity with patient-reported well-being 3 months after a symptomatic illness. *JAMA Netw Open* 2022;5(12):e2244486. <https://doi.org/10.1001/jamanetworkopen.2022.44486>.
- Cutler DM. *The Economic Cost of Long COVID: An Update*. Available at: https://scholar.harvard.edu/files/cutler/files/long_covid_update_7-22.pdf. Accessed January 4, 2022.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.amjmed.2023.01.005>.

Supplementary Table 1 Frequencies and Percentages, Unless Otherwise Indicated, of Characteristics of Long COVID Studies in High-Impact Journals

	All studies (N = 83)	Observational studies without a comparison group or control arm (N = 35)	Studies with a comparison group or control arm (N = 48; including the 3 randomized studies)	Randomized studies (N = 3)	P value comparing studies with and without a comparison group
Year that enrollment completed					0.50
2019	1 (1.2)	0	1 (2.1)	0	
2020	54 (65.1)	26 (74.3)	28 (58.3)	0	
2021	21 (25.3)	5 (14.3)	16 (33.3)	0	
2022	2 (2.4)	1 (2.9)	1 (2.1)	3 (100)	
Not indicated	5 (6.0)	3 (8.6)	2 (4.2)	0	
Journal impact factor, median (IQR)	18 (12, 21)	18 (12, 21)	17 (12, 24)	34 (27, 49)	
Altmetric score	173 (35, 959)	79 (32, 959)	206 (44, 930)	602 (338, 720)	.52
Geographic region					.26
Europe	57 (68.7)	25 (71.4)	32 (66.7)	2 (66.7)	
US	16 (19.3)	4 (11.4)	12 (25.0)	1 (33.3)	
Other	9 (10.8)	5 (14.3)	4 (8.3)	0	
Multiple	1 (1.2)	1 (2.9)	0	0	
Study design					.007
Case control	5 (6.0)	0	5 (10.4)	0	
Cross-sectional	15 (18.1)	9 (25.7)	6 (12.5)	0	
Prospective cohort	46 (55.4)	24 (68.6)	22 (45.8)	0	
Randomized clinical trial	3 (3.6)	0	3 (6.2)	3 (100)	
Retrospective observational	14 (16.9)	2 (5.7)	12 (25.0)	0	
Population					.04
Community (general)	12 (14.5)	8 (22.9)	4 (8.3)	1 (33.3)	
Health care workers	1 (1.2)	0	1 (2.1)	0	
Inpatients	17 (20.5)	11 (31.4)	6 (12.5)	0	
Outpatients	33 (39.8)	10 (28.6)	23 (47.9)	1 (33.3)	
In- and outpatients	10 (12.0)	4 (11.4)	6 (12.5)	0	
National database	5 (6.0)	0	5 (10.4)	0	
Social media	3 (3.6)	2 (5.7)	1 (2.1)	0	
Not indicated	2 (2.4)	0	2 (4.2)	1 (33.3)	
Total number of study participants	599 (150, 7584)	797 (270, 2185)	400 (87, 51442)	150 (102, 216)	.61
Age, median (IQR)*	49 (45, 56)	50 (47, 57)	48 (43, 55)	47 (46, 48)	.13
Age category					.38
Adult	50 (61.0)	22 (64.7)	28 (58.3)	2 (66.7)	
Child	4 (4.9)	0	4 (8.3)	0	
Both adults and children	13 (15.9)	6 (17.6)	7 (14.6)	0	
Not indicated	15 (18.3)	6 (17.6)	9 (18.8)	1 (33.3)	
Vaccination accounted for					.32
Yes (postvaccine)	13 (15.7)	4 (11.4)	9 (18.8)	0	
No (postvaccine)	8 (9.6)	3 (8.6)	5 (10.4)	1 (33.3)	
No (prevaccine)	53 (63.9)	26 (74.3)	27 (56.2)	0	
No (youth)	4 (4.8)	0	4 (8.3)	0	
Not indicated	5 (6.0)	2 (5.7)	3 (6.2)	2 (66.7)	

COVID = coronavirus disease; IQR = interquartile range.

*The number of missing for age was 12.

Supplementary Table 2 Characteristics of Long COVID Definitions and Strategies to Adjust for Confounding in the Study Design and Analysis in Studies Examining Long COVID Published in High-Impact Journals

	All studies (N = 83)	Observational studies without a comparison group or control arm (N = 35)	Studies with a comparison group or control arm (N = 48; includes the 3 randomized studies)	Randomized studies (N = 3)	P value comparing studies with and without a comparison group
Long COVID definitions					
Duration of COVID symptoms					.12
3-4 weeks	23 (27.7)	6 (17.1)	17 (35.4)	1 (33.3)	
6 weeks	3 (3.6)	1 (2.9)	2 (4.2)	0	
8 weeks	10 (12.0)	7 (20.0)	3 (6.2)	0	
10-12 weeks	21 (25.3)	9 (25.7)	12 (25.0)	1 (33.3)	
16 or more weeks	3 (3.6)	0	3 (6.2)	0	
Not indicated	23 (27.7)	12 (34.3)	11 (22.9)	1 (3.3)	
COVID diagnosis					.16
Confirmed	58 (69.9)	23 (65.7)	35 (72.9)	1 (33.3)	
Unconfirmed	19 (22.9)	11 (31.4)	8 (16.7)	2 (66.7)	
Not indicated	6 (7.2)	1 (2.9)	5 (10.4)	0	
Has a specific long COVID definition	44 (53.0)	24 (68.6)	20 (41.7)	2 (66.7)	.03
Number of long COVID symptoms, median (IQR)*	21 (13, 30)	21 (19, 39)	18 (12, 23)	32 (28, 35)	.05
Funding					.85
Industry	4 (4.8)	1 (2.9)	3 (6.2)	1 (33.3)	
Other	58 (69.9)	26 (74.3)	32 (66.7)	2 (66.7%)	
None	8 (9.6)	3 (8.6)	5 (10.4)	0	
Not indicated	13 (15.7)	5 (14.3)	8 (16.7)	0	
Author conflict of interest					.37
Industry	39 (47.0)	15 (42.9)	24 (50.0)	1 (33.3)	
Other	4 (4.8)	1 (2.9)	3 (6.2)	0	
None	38 (45.8)	19 (54.3)	19 (39.6)	2 (66.7)	
Not indicated	2 (2.4)	0	2 (4.2)	0	
Reported prevalence of long COVID	32 (38.6)	0	0	0	
Prevalence of long COVID, median (IQR)†	48 (36, 66)	51 (42, 68)	46 (14, 65)	NA	.18
Prevalence of long COVID in the comparison group, median (IQR)†	34 (7, 38)	NA	34 (7, 38)	NA	
Study assessment type					.09
Prevention	2 (2.4)	2 (5.7)	0	0	
Risk factors, in addition to symptoms	24 (28.9)	11 (31.4)	13 (27.1)	0	
Symptoms	52 (62.7)	22 (62.9)	30 (62.5)	0	
Treatment	5 (6.0)	0	5 (10.4)	3 (100%)	
Comparison group and adjustment					
Adjustment in analysis	42 (50.6)	16 (45.7)	26 (54.2)	1	.59
Number of variables adjusted					.23
1	3 (3.6)	0	3 (6.2)	0	
2	8 (9.6)	2 (5.7)	6 (12.5)	0	
3	4 (4.8)	3 (8.6)	1 (2.1)	0	
4	3 (3.6)	2 (5.7)	1 (2.1)	0	
5 or more	21 (25.3)	9 (25.7)	12 (25.0)	0	
Unspecified	3 (3.6)	0	3 (6.2)	1 (33.3)	
Matching	22 (26.5)	NA	22 (45.8)	0	
Used a negative control outcome	4 (4.8)	NA	4 (8.3)	0	
Used a positive control outcome	3 (3.6)	NA	3 (6.2)	0	

Supplementary Table 2 (Continued)

	All studies (N = 83)	Observational studies without a comparison group or control arm (N = 35)	Studies with a comparison group or control arm (N = 48; includes the 3 randomized studies)	Randomized studies (N = 3)	P value comparing studies with and without a comparison group
Long COVID definitions					
Used a negative exposure outcome	1 (1.2)	NA	1 (2.1)	0	
Comparison of groups at baseline					
Yes	23 (27.7)	NA	23 (47.9)	0	
Few (demographics plus limited health status variables)	10 (12.0)	NA	10 (20.8)	1 (33.3)	
Demographics only	6 (7.2)	NA	6 (12.5)	1 (33.3)	
No comparison of baseline variables	44 (53.0)	NA	9 (18.8)	1 (33.3)	
Comparison groups					
COVID vs COVID	4 (4.8)	NA	4 (8.3)	0	
COVID vs no COVID	27 (32.5)	NA	27 (56.2)	0	
COVID vs other	3 (3.6)	NA	3 (6.2)	0	
Long COVID vs COVID	9 (10.8)	NA	9 (18.8)	0	
Long COVID vs long COVID	4 (4.8)	NA	4 (8.3)	3 (100)	
Long COVID vs no COVID	1 (1.2)	NA	1 (2.1)	0	
None	35 (42.2)	35 (100)	0	0	

COVID = coronavirus disease; IQR = interquartile range; NA = not applicable

*The number of missing for number of long COVID symptoms was 39.

†There were 33 observations used in determining long COVID prevalence in the exposure group and 8 observations used in determining long COVID prevalence in the comparator group.