

Risk Predictors and Symptom Features of Long COVID Within a Broad Primary Care Patient Population Including Both Tested and Untested Patients

Rupert Jones,^{1,3} Andrew Davis,^{1,2} Brooklyn Stanley,^{1,2} Steven Julious,⁴ Dermot Ryan,⁵ David J Jackson,⁶ David MG Halpin,⁷ Katherine Hickman,⁸ Hilary Pinnock,⁹ Jennifer K Quint,¹⁰ Kamlesh Khunti,¹¹ Liam G Heaney,¹² Phillip Oliver,¹³ Salman Siddiqui,¹³ Ian Pavord,¹⁴ David HM Jones,¹⁵ Michael Hyland,^{3,16} Lewis Ritchie,¹⁷ Pam Young,¹⁸ Tony Megaw,¹⁸ Steve Davis,¹⁹ Samantha Walker,²⁰ Stephen Holgate,²¹ Sue Beecroft,²² Anu Kemppinen,^{1,2} Francis Appiagyei,^{1,2} Emma-Jane Roberts,^{1,2} Megan Preston,^{1,2} Antony Hardjojo,^{1,2} Victoria Carter,^{1,2} Marije van Melle,^{1,2} David Price^{1,2,17}

¹Optimum Patient Care, Cambridge, UK; ²Observational and Pragmatic Research Institute, Singapore, Singapore; ³Faculty of Health, University of Plymouth, Plymouth, UK; ⁴University of Sheffield, South Yorkshire, UK; ⁵Usher Institute, University of Edinburgh, Edinburgh, UK; ⁶Guy's & St Thomas' NHS Trust, School of Immunology & Microbial Sciences, King's College London, London, UK; ⁷University of Exeter Medical School, College of Medicine and Health, University of Exeter, Exeter, UK; ⁸Bradford and Leeds Clinical Commissioning Group, Leeds, UK; ⁹Asthma UK Centre for Applied Research, Usher Institute, The University of Edinburgh, Edinburgh, UK; ¹⁰National Heart & Lung Institute, Imperial College London, London, UK; ¹¹Diabetes Research Centre, University of Leicester, Leicester, UK; ¹²Wellcome-Wolfson Centre for Experimental Medicine, Queen's University Belfast, Belfast, Northern Ireland; ¹³Institute for Lung Health, Leicester National Institute for Health Research Biomedical Research Centre, University of Leicester, Leicester, UK; ¹⁴Respiratory Medicine Unit and Oxford Respiratory NIHR BRC, Nuffield Department of Medicine, University of Oxford, Oxford, UK; ¹⁵Box Surgery, Wiltshire, UK; ¹⁶Plymouth Marjon University, Plymouth, UK; ¹⁷Centre of Academic Primary Care, Division of Applied Health Sciences, University of Aberdeen, Aberdeen, UK; ¹⁸Wellbeing Software, Mansfield, UK; ¹⁹Interface Clinical Services, Leeds, UK; ²⁰Asthma UK & British Lung Foundation, London, UK; ²¹Clinical and Experimental Sciences, University of Southampton, Southampton, UK; ²²OPEN Health, Buckinghamshire, UK

Correspondence: David Price
Academic Primary Care, Division of Applied Health Sciences, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, UK
Tel +65 3105 1489
Email dprice@opri.scg

Introduction: Symptoms may persist after the initial phases of COVID-19 infection, a phenomenon termed long COVID. Current knowledge on long COVID has been mostly derived from test-confirmed and hospitalized COVID-19 patients. Data are required on the burden and predictors of long COVID in a broader patient group, which includes both tested and untested COVID-19 patients in primary care.

Methods: This is an observational study using data from Platform C19, a quality improvement program-derived research database linking primary care electronic health record data (EHR) with patient-reported questionnaire information. Participating general practices invited consenting patients aged 18–85 to complete an online questionnaire since 7th August 2020. COVID-19 self-diagnosis, clinician-diagnosis, testing, and the presence and duration of symptoms were assessed via the questionnaire. Patients were considered present with long COVID if they reported symptoms lasting ≥ 4 weeks. EHR and questionnaire data up till 22nd January 2021 were extracted for analysis. Multivariable regression analyses were conducted comparing demographics, clinical characteristics, and presence of symptoms between patients with long COVID and patients with shorter symptom duration.

Results: Long COVID was present in 310/3151 (9.8%) patients with self-diagnosed, clinician-diagnosed, or test-confirmed COVID-19. Only 106/310 (34.2%) long COVID patients had test-confirmed COVID-19. Risk predictors of long COVID were age ≥ 40 years (adjusted Odds Ratio [AdjOR]=1.49 [1.05–2.17]), female sex (adjOR=1.37 [1.02–1.85]), frailty (adjOR=2.39 [1.29–4.27]), visit to A&E (adjOR=4.28 [2.31–7.78]), and hospital admission for COVID-19 symptoms (adjOR=3.22 [1.77–5.79]). Aches and pain (adjOR=1.70 [1.21–2.39]), appetite loss (adjOR=3.15 [1.78–5.92]), confusion and disorientation (adjOR=2.17 [1.57–2.99]), diarrhea (adjOR=1.4 [1.03–1.89]), and persistent dry cough (adjOR=2.77 [1.94–3.98]) were symptom features statistically more common in long COVID.

Conclusion: This study reports the factors and symptom features predicting long COVID in a broad primary care population, including both test-confirmed and the previously missed group of COVID-19 patients.

Keywords: SARS-CoV-2, questionnaire, observational study, frailty, chronic diseases

Plain Language Summary

Why was the study done?

Symptoms of COVID-19 may continue after the initial phases of the infection into a condition termed long COVID. Most research on long COVID has focused on patients

who had received test confirmation for COVID-19 or were hospitalized with the condition. However, many COVID-19 cases went undetected due to the shortage of testing in the early phase of the pandemic. This study aims to assess the prevalence of long COVID, and the patient characteristics and symptoms which can predict the development of long COVID in a broader primary care population, including both COVID-19 patients identified via testing and those who did not receive test confirmation.

What did the researchers do and find?

This study analyzed the data collected within platform C19, a research database combining data from routine primary care visits and patient-reported information from a questionnaire. Patients were considered to have long COVID if they reported having symptoms of COVID-19 lasting at least 4 weeks. Factors that predicted the development of long COVID were female sex, age above 40, being frail, history of a visit to accidents and emergency, and admission to the hospital for COVID-19 infection. The presence of aches and pains, loss of appetite, confusion and disorientation, diarrhea, and persistent dry cough indicates patients whose symptoms will likely develop long COVID.

What do these results mean?

This study provides new insights on long COVID in a broad range of patients in primary care in the UK. These results may assist healthcare providers to identify and manage COVID-19 patients at risk of persisting symptoms.

Introduction

There have been multiple reports of symptoms of COVID-19 infection persisting long after the initial phases of the infection, a phenomenon termed long COVID or post-acute COVID-19.^{1,2} There is no single agreed definition for long COVID. The UK National Institute of Health and Care Excellence (NICE)^{2,3} and the Wellcome Trust foundation⁴ defined long COVID as symptoms lasting 4 weeks or more since onset, while other definitions include a duration cut-off of 12 weeks.⁵

Long COVID appears to be common, persistent, and debilitating, causing a wide range of physiological and cognitive disabilities irrespective of patient age and initial disease severity.^{1,6-9} The impact on social well-being¹⁰ highlights the importance of addressing the persisting symptoms of COVID-19 infection. The common presentation of persisting symptoms includes fatigue, breathlessness, chest pain, and body aches.^{9,11,12} While comparison is difficult due to the lack of a uniform definition, especially in the early days of the pandemic, a review by Nalbandian et al suggested up to 87.4% of COVID-19 patients may remain symptomatic months since the initial symptom presentation.¹ Symptom of COVID-19 infection

has also been reported to persist one year post-hospital discharge.¹³

Reports of prevalence from a more general patient population suggest a much lower prevalence. The UK COVID-19 Infection Survey produced experimental results showing a national estimate of one in five patients who were tested positive for COVID-19 were symptomatic 5 weeks into the infection, with a further one in ten patients' symptoms lasted beyond 12 weeks as of 16th December 2020.¹⁴ In the US, telephone interviews of RT-PCR-confirmed symptomatic outpatient adults reported 35% having not recovered after 2-3 weeks, and only 1 in 5 young healthy adults.¹¹

A better understanding of the symptom characteristics and the risk predictors for long COVID is needed to identify patients at risk of developing the condition. However, current data on persisting symptoms of COVID-19 has been mostly derived from cases identified via positive test results.^{7-11,14-16} While the exact numbers are not known, estimates suggested that the majority of COVID-19 cases went undetected,^{17,18} likely due to limited testing capacity in the early phase of the pandemic. Furthermore, many studies on the long-term symptoms of COVID-19 have focused primarily on hospitalized patients.^{7-9,12,19-23} The full burden of long COVID and factors which predict the risk of developing long COVID have not been clearly defined in the primary care population including the untested patients who may have a milder initial disease but may still be at risk of developing long COVID.²⁴ Supplementing electronic health data (EHR) data with patient-reported information can provide additional insight on these previously missed patients with symptoms of COVID-19 managed in primary care.

In response to the pandemic, Optimum Patient Care (OPC) UK established the COVID-19 Quality Improvement (QI) program to help practices identify and manage patients affected by COVID-19 and related problems including pre-existing chronic diseases. (<https://optimumpatientcare.org/covid-qi/>). As part of the program, participating general practitioners invite their patients to complete an online questionnaire covering questions on demographics, co-morbid conditions, COVID-19 status, symptoms, and testing. Platform C19 systematically assimilates extensive EHR data with patient-reported outcomes (PROs) through questionnaires in the primary care population, creating a unique platform that enables new analyses and insights on the pandemic.

This study aims to investigate the prevalence, symptom features, and demographical and clinical factors which predict the development of long COVID within a broader primary care population within Platform C19 which

includes both test-confirmed COVID-19 patients and COVID-19 patients who did not receive test confirmation.

Methods

Study Design and Data Source

This is an observational study using EHR data and patient-reported information data stored in Platform C19. EHR data is derived from the OPCR, a de-identified primary care database holding records for more than 12 million patients from over 800 GP practices across the UK. The OPCR integrates with all UK clinical systems (EMIS, TPP SystemOne, InPS Vision, Microtest Evolution) and holds a long retrospective period of data extending from the conception of summary diagnostic data collection (median time in the database of 13 years). The OPCR data is further enriched with routine EHR data linked with patient-reported information from GP practices participating in OPC QI programs. The questionnaire was designed via a consensus of experts from the Platform C19 steering committee.

Patients aged 18–85 years at the start of the pandemic (1st March 2020) from practices participating in the COVID-19 QI program were invited by text message from the GP to complete an online COVID-19 questionnaire as part of the QI starting from 7th August 2020. Patients that have opted out from data-sharing for research²⁵ or from receiving text messages were excluded.

De-identified EHR and questionnaire data collected from 7th August 2020 up till 22nd January 2021 from all patients who responded were included. To prepare data for analysis, questionnaire responses and relevant electronic medical records were collated, cleaned, and summarized. If patients had multiple questionnaires on record only the most recent or most complete one was used. Data is stored in a secure, enterprise database running SQL-Server (Windows Server 2019 Standard (10.0) Version 15.0.2080.9).

Study Variables

COVID-19 self-diagnosis, clinical diagnosis, and testing were ascertained via the questionnaire. Patients were asked whether they believed they have had a COVID-19 infection, whether they had been diagnosed with COVID-19 by a healthcare professional, and whether they had been tested for COVID-19 and the result of the test. Patients with COVID-19 status were split into three categories: self-diagnosed but not diagnosed by a clinician or test-

confirmed, diagnosed by a clinician but not test-confirmed, and test-confirmed positive.

The demographic variables extracted from the EHR include age, sex, and smoking status. Data on ethnicity and BMI were supplemented by the questionnaire. For analysis, age was split into an 18–40 year group and an age 40+ group. BMI was split into below 30 kg/m² and greater than or equal to 30 kg/m²; the threshold for obesity. Smokers were classified as non-smokers, ex-smokers, current smokers, or having an unknown smoking status.

The presence of underlying chronic co-morbid diseases was extracted from both EHR and questionnaire. Diseases investigated were asthma, COPD, diabetes, heart disease or heart failure, kidney disease, and any combination of these diseases. Frailty was defined as either recorded frailty in the EHR or responding to be frail (“you have medical problems that limit how active you are, and you need help with daily activities and personal care”) in the question assessing the level of fitness. Patients were also asked if they visited the accident and emergency department, if they were admitted to a hospital, and if they were admitted to intensive or critical care for their COVID-19 infection or symptoms. Presence, start date, and end date of symptoms typical of COVID-19 since January 2020 were asked in the questionnaire. The list of questions within the questionnaire used for analysis and the list of COVID-19 symptoms are available in the [Supplementary Material](#).

Outcome Definition

The primary outcome of this study is the presence of long COVID based on COVID-19 status and symptom duration based on symptom start and end date as reported by the patients. Patients were considered to have had long COVID if they were self-diagnosed, clinician-diagnosed, or test-confirmed for COVID-19 and have symptoms of COVID which lasted for more than 4 weeks based on the NICE guideline definition for long COVID.³

Statistical Analysis

All analyses were conducted using Rv4.0.3 (R Core team, 2020). Univariable logistic regression was conducted to summarize each demographic and clinical variable individually. Multivariable logistic regression models were created using all demographic variables, hospital visits for COVID-19, frailty, chronic co-morbid conditions, and COVID-19 status as predictors and long COVID as the response variable in comparison with COVID-19 with <4

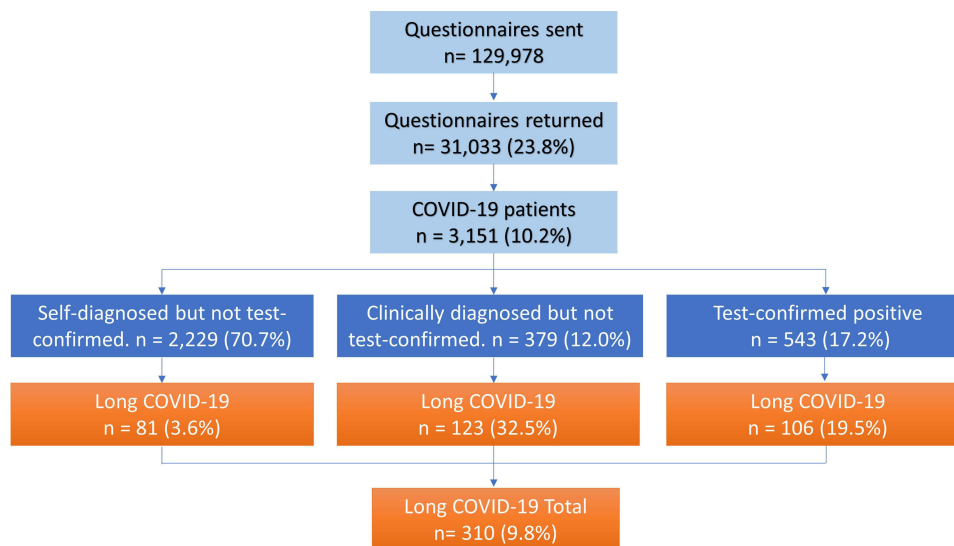


Figure 1 Patient flow.

week symptom duration. Two separate regression models were created, one with each underlying co-morbid chronic disease entered as separate variables (split condition model) and another with the presence of any underlying co-morbid chronic disease combined as a single variable (combined condition model).

Univariable and multivariable logistic regression analyses to compare the symptom features of long COVID and COVID-19 with <4 week symptom duration were also conducted. The multivariable regression model for symptom comparison included every symptom adjusted for factors that significantly predict long COVID from the analysis above. Odds ratios were estimated along with 95% confidence intervals to evaluate variable influence. There is no allowance for multiplicity. Sensitivity analyses were also conducted on long COVID using 12 week symptom duration compared to COVID-19 with symptoms lasting <12 weeks.

Results

Patient Breakdown

The patient flowchart is presented in [Figure 1](#). Questionnaires were sent to 129,978 eligible patients, of whom 31,033 (23.9%) responded. COVID-19 was reported by 3151/31,033 (10.2%) respondents, of whom 310/3151 (9.8%) were present with long COVID. Of those, 81/310 (26.1%) were self-diagnosed without clinician diagnosis or test confirmation, 123/310 (39.7%) received a diagnosis by a clinician but not test confirmation, and 106/310 (34.2%) had test-confirmed COVID-19.

The demographic and clinical characteristics of the patients are presented in [Table 1](#).

Univariable Analysis for Risk Predictors for Long COVID

A statistically higher proportion of patients who were clinician-diagnosed but not test-confirmed had long COVID than those who were test-confirmed (39.7% vs 34.2%; OR = 1.98 [1.47–2.68]) ([Supplementary Table 1](#)). Conversely, statistically lower proportion of patients who self-diagnosed with neither clinician diagnosis nor testing had long COVID compared to test-confirmed patients (26.1% vs 34.2%; OR = 0.15 [0.11–0.21]).

Patients with long COVID compared to, by all definitions of self-diagnosed, clinically diagnosed and test-confirmed, were statistically more likely to be frail (6.8% vs 2.8%; OR = 2.51 [1.49–4.04]), be above the age 40 (83.2% vs 76.7%; OR = 1.51 [1.12–2.08]), and female (72.3% vs 64.2%; OR = 1.45 [1.12–1.89]) compared to patients with <4 weeks of symptoms. Of the chronic co-morbid diseases, COPD, bronchitis or emphysema (7.4% vs 2.5%; OR = 3.13 [1.89–5.00]) and heart failure or heart disease (9.0% vs 5.0%; OR = 1.89 [1.21–2.84]) were statistically more prevalent in patients with long COVID. Statistically more patients with long COVID had visited an accident and emergency department (10.6% vs 2.2%; OR = 5.29 [3.11–8.77]), had been admitted to a hospital (9.0% vs 1.9%; OR = 5.79 [3.38–9.70]), and had been admitted to an intensive or critical care unit (1.6% vs 0.5%; OR = 4.07 [1.30–10.88]) for their COVID-19 infection or symptoms.

Table 1 Patient Demographics and Clinical Characteristics

Variable ^a		Long COVID n=310	<4 Weeks COVID-19 n=2841
Age (years)	Mean (SD)	52.1 (12.2)	51.4 (39.5)
	Median (IQR)	53 (44–60)	52 (40–61)
	18 – <40	52 (16.8%)	662 (23.3%)
	≥40	258 (83.2%)	2179 (76.7%)
Sex	Female	224 (72.3%)	1825 (64.2%)
BMI (kg/m ²)	<18.5	3 (1%)	33 (1.2%)
	18.5–24.9	63 (21.4%)	679 (25.7%)
	25–29.9	92 (31.3%)	837 (31.7%)
	30–39.9	102 (34.7%)	874 (33.1%)
	≥40	34 (11.6%)	219 (8.3%)
Smoking status	Current smokers	38 (12.5%)	340 (12.2%)
	Ex-smokers	132 (43.3%)	1129 (40.6%)
	Never smoked	135 (44.3%)	1310 (47.1%)
Ethnicity	White	278 (96.2%)	2162 (91.8%)
	Non-White	11 (3.8%)	194 (8.2%)
Frailty	No reported	289 (93.2%)	2761 (97.2%)
	Reported Frailty	21 (6.8%)	80 (2.8%)
Hospital visits for COVID-19 symptoms or infection	Any	51 (16.5%)	101 (3.6%)
	Visited A&E	33 (10.6%)	63 (2.2%)
	Admitted to hospital	28 (9.0%)	54 (1.9%)
	Admitted to intensive Care	5 (1.6%)	13 (0.5%)
Co-morbid diseases	CKD	11 (3.5%)	93 (3.3%)
	Hypertension	60 (19.4%)	461 (16.2%)
	Heart failure	5 (1.6%)	22 (0.8%)
	Other heart diseases	26 (8.4%)	137 (4.8%)
	Lung Cancer	0 (0%)	0 (0%)
	COPD, bronchitis or emphysema	23 (7.4%)	71 (2.5%)
	Asthma (ever)	72 (23.2%)	566 (19.9%)
	Active Asthma	53 (17.1%)	475 (16.7%)
	Diabetes	78 (25.2%)	588 (20.7%)
	Depression or anxiety ^b	84 (27.1%)	559 (19.7%)

Notes: Patients were considered to have COVID-19 if they were self-diagnosed, clinician-diagnosed, or test-confirmed for COVID-19. Patients presented with long COVID also have symptoms lasting at least 4 weeks. ^aNumbers are presented as n (%) of valid responses unless specified. ^bPrior to January 2020.

Abbreviations: CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease.

Sensitivity analysis using 12 week cut-off definition showed similar results, other than for admission to intensive or critical care for COVID-19 symptoms ([Supplementary Table 2](#)).

Multivariable Analysis for Predictors for Long COVID

Following multivariable adjustment, patients who were only self-diagnosed without clinician diagnosis or test-confirmation were statistically less likely to have long COVID (adjusted OR

= 0.16 [0.11–0.22]) than patients with test-confirmed COVID-19. Patients who were clinician-diagnosed without test-confirmation also remained statistically more at-risk of long COVID (adjusted OR = 1.99 [1.44–2.75]) ([Table 2](#)).

Frailty (adjusted OR = 2.39 [1.29–4.27]), age above 40 years (adjusted OR = 1.49 [1.05–2.17]), and female sex (adjusted OR = 1.37 [1.02–1.85]) were independently associated with long COVID. However, admission to intensive or critical care and all co-morbid chronic diseases were no longer statistically significant. Similar results were obtained using the

Table 2 Split Condition Multivariable Model for Risk Predictors for Long COVID

Variable		Adjusted OR (95% CI) ^a
COVID-19 status	Test confirmed positive	Reference
	Clinical diagnosis without test confirmation	1.99 (1.44–2.75) ^b
	Self-diagnosed without test confirmation	0.16 (0.11–0.22) ^b
Hospital visit for COVID-19 symptoms or infection	None	Reference
	Visited A & E	4.28 (2.31–7.78) ^b
	Admitted to hospital	3.22 (1.77–5.79) ^b
	Admitted to intensive or critical care	1.29 (0.39–3.71)
Underlying chronic co-morbid diseases	No disease	Reference
	Asthma	1.06 (0.76–1.44)
	COPD, bronchitis or emphysema	1.53 (0.84–2.72)
	Diabetes	1.07 (0.78–1.45)
	Ischemic heart disease or heart failure	1.26 (0.75–2.05)
	Kidney disease	1.22 (0.56–2.43)
Age	18–40 years	Reference
	≥40 years	1.49 (1.05–2.17) ^b
Sex	Male	Reference
	Female	1.37 (1.02–1.85) ^b
BMI	<30 kg/m ²	Reference
	≥30 kg/m ²	1.05 (0.80–1.37)
Smoking status	Non-Smoker	Reference
	Ex-Smoker	1.15 (0.86–1.53)
	Current Smoker	1.43 (0.91–2.20)
	Unknown	0.73 (0.11–2.55)
Frailty	No reported frailty	Reference
	Reported frailty	2.39 (1.29–4.27) ^b

Notes: ^aAdjusted for all demographic variables, hospital visits for COVID-19, frailty, chronic co-morbid conditions, and COVID-19 status; ^bStatistically significant at $p < 0.005$.

combined disease multivariable model ([Supplementary Table 3](#)).

Sensitivity analysis using the 12 weeks cut-off definition showed similar results except for ischemic heart failure or heart disease (adjusted OR = 1.72 [1.00–2.87]) in the split condition model ([Supplementary Tables 4 and 5](#)).

Symptom Features of Long COVID

The presence of symptoms in patients with long COVID and COVID-19 with shorter symptom duration is presented in [Figure 2](#). Almost every (95.2%) patients with long COVID reported a loss of appetite. Other commonly reported symptoms among long COVID patients were persistent dry cough (80.3%), chest pain (79.7%), and fatigue or tiredness (72.9%). The symptom features of long COVID using the 12 week cut-off definition showed a similar pattern ([Supplementary Figure 1](#)).

Appetite loss (adjusted OR = 3.15 [1.78–5.92]), persistent dry cough (adjusted OR = 2.77 [1.94–3.98]), confusion or disorientation (adjusted OR = 2.17 [1.57–2.99]), aches and pain (adjusted OR = 1.70 [1.21–2.39]), and diarrhea (adjusted OR = 1.40 [1.03–1.89]) were statistically more common in patients with long COVID ([Table 3](#)). Surprisingly, shortness of breath was statistically less common (adjusted OR = 0.68 [0.50–0.91]).

The results using the 12 weeks cut-off were similar except “fatigue and tiredness” was also statistically less common in long COVID patients (adjusted OR = 0.66 [0.46–0.96]) ([Table 3](#)).

Discussion

Prevalence of Long COVID

In this study, we identified long COVID, defined as COVID-19 with symptoms lasting at least 4 weeks, in nearly one-tenth of

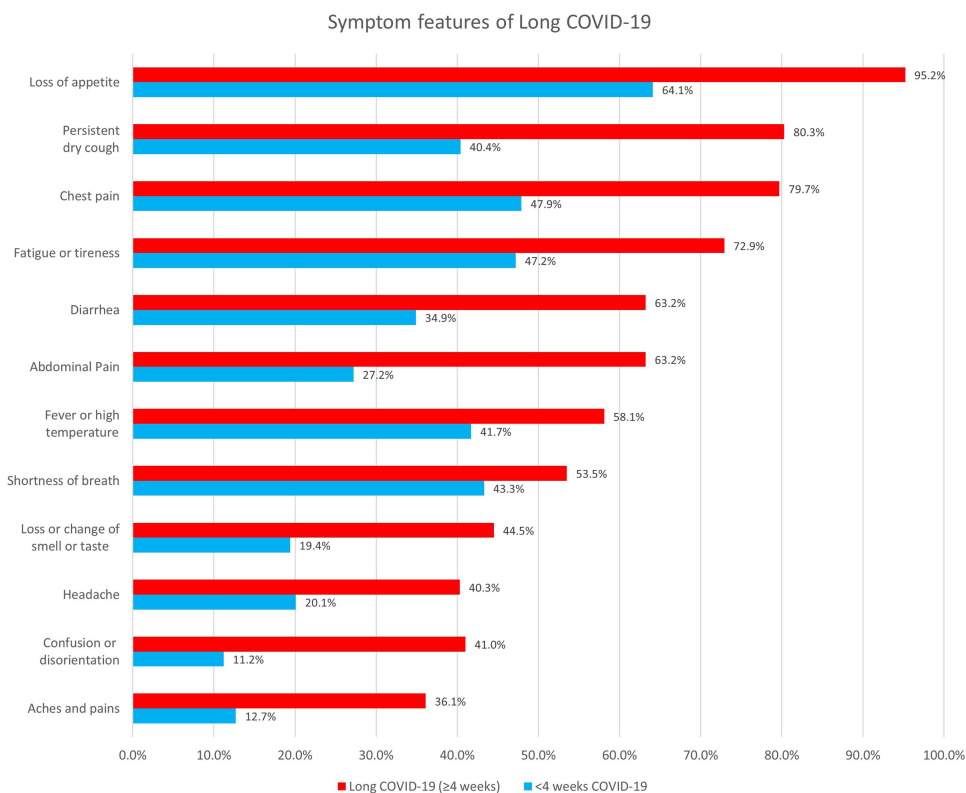


Figure 2 Presence of symptoms in COVID-19 patients.

COVID-19 cases. This is much lower than the rates reported in other studies which have been reported to be more than half the patient population.^{13,22} However, this could be due to the difference in the method of symptom assessment and patient population. The prevalence observed in this study is closer, albeit still lower, compared to more similar general populations from the UK COVID-19 infection survey and the proportion reported in an observational study based on users of the COVID Symptom Study mobile application in the UK (13.3%).^{14,16} As the patients in the current study were not limited to test confirmed COVID-19 cases, our results may be reflective of the prevalence of long COVID in a more general COVID-19 population. Indeed, only slightly above one-third of long COVID cases identified in this study were test-confirmed, highlighting the magnitude of the important yet previously unquantified population of long COVID patients.

Predictors of Long COVID

Patients who were diagnosed by a clinician but never test-confirmed were more likely to have long COVID than those who were test-confirmed positive. We speculate that this may be due to patients with milder symptoms only received examination and testing after their symptoms have persisted. Patients who did not receive test

confirmation may also make up the group of patients from the initial phases of the pandemic when testing and awareness of long COVID were limited. In contrast, significantly fewer self-diagnosed patients who believed themselves to be infected, without a clinician diagnosis or test confirmation, had long COVID. This may be as patients whose COVID-19 symptoms resolved rapidly did not feel the need to be checked by a clinician or tested. Alternatively, a proportion of these patients may not have had COVID-19 infection and this is difficult to assess.

Older age and female sex were independently associated with long COVID. Sexual differences in COVID-19 related symptomatology and mortality had been widely recognized. While male sex had been reported to be associated with a higher risk of COVID-19 hospitalization and mortality,²⁶ studies have reported female sex to be at a higher risk for prolonged COVID-19 symptoms.^{7,8,16,22,23} Biological differences could contribute to the difference in symptom persisting between sex, though gender differences in symptom reporting behavior may have also played a part.²⁷

Frailty significantly predicted long COVID even after adjustment with age and chronic diseases. Frailty was previously reported in a multicenter cohort of hospitalized elderly patients to be a strong predictor for poorer outcomes of

Table 3 Multivariable Analysis for Symptom Features Indicative of Long COVID

Symptoms Since January 2020	Adjusted OR (95% CI) ^a	
	4 Weeks Definition	12 Weeks Definition
Abdominal pain	1.33 (0.97–1.83)	1.39 (0.99–1.97)
Aches and pain	1.70 (1.21–2.39) ^b	1.49 (1.02–2.16) ^b
Appetite loss	3.15 (1.78–5.92) ^b	2.91 (1.55–5.91) ^b
Chest pain	1.42 (0.99–2.04)	1.42 (0.96–2.13)
Confusion or disorientation	2.17 (1.57–2.99) ^b	2.14 (1.52–3.03) ^b
Diarrhea	1.40 (1.03–1.89) ^b	1.42 (1.02–1.98) ^b
Fatigue or tiredness	0.83 (0.59–1.17)	0.66 (0.46–0.96) ^b
Fever	0.78 (0.58–1.05)	0.88 (0.63–1.22)
Headache	0.98 (0.71–1.34)	0.99 (0.70–1.40)
Loss/change of smell or taste	1.16 (0.84–1.58)	1.13 (0.80–1.58)
Persistent dry cough	2.77 (1.94–3.98) ^b	3.24 (2.17–4.92) ^b
Shortness of breath	0.68 (0.50–0.91) ^b	0.69 (0.50–0.95) ^b

Notes: ^aIn comparison with COVID-19 of shorter symptom duration and adjusted with COVID-19 status, Age, Sex, Frailty, and Hospital Visits. ^bStatistically significant at $p < 0.005$.

COVID-19 infection.²⁸ The National Institute for Health and Care Excellence (NICE) of the UK advised frailty to be included as part of the assessment for patients admitted into intensive care during the pandemic.²⁹ Awareness of frailty as a risk factor for long COVID may be useful for clinicians when patients present with new-onset prolonged diverse symptoms as reported in the present study.

Patients who visited an emergency department and who were admitted to the hospital for COVID-19 symptoms or infections were statistically more likely to have long COVID. Similar to our observation on COVID-19 status, this may be due to the increasing likeliness of patients visiting the hospital as their symptoms persisted or due to the higher severity of their illness. Admission to intensive or critical care was no longer significant following multivariable adjustment, and this may have been due to the small numbers.

Symptom Features of Long COVID

Multiple symptoms across different organ systems were common among our patients, consistent with known information on COVID-19 symptomatology.¹ Loss of appetite, persistent chest pain, persistent dry cough, and fatigue or tiredness were the most commonly reported symptoms. The range of symptoms observed in this study is similar to that observed by previous studies which commonly list symptoms of fatigue, body aches,

breathlessness, and cough.^{9,11,12} However, this study further compares patients with long COVID to patients with shorter symptom duration to investigate the symptoms that can potentially predict if a patient will develop long COVID.

Individually, all symptoms were significantly more commonly reported by long COVID patients than patients with shorter symptom duration. This is similar to the observation in the study using the COVID Symptom Study application.¹⁶ Following multivariable adjustment of the symptoms with each other, several symptoms were no longer independently associated with long COVID. Additionally, shortness of breath and fatigue or tiredness were found to have an inverse association with long COVID using the 12 weeks cut-off definition. These results may indicate that these symptoms were pervasive enough in the acute phase of the infection that they were not strong indicators of COVID-19 infection which will persist into long COVID. Symptoms of aches and pain, appetite loss, confusion and disorientation, diarrhea, and persistent dry cough were the symptom features that predicted long COVID within our patient population. Diarrhea, but not fever or dyspnea, has also been previously reported to be associated with persistent symptoms among hospitalized patients in India.³⁰

Strengths

To our knowledge, this study is the first to identify and report on the scale of long COVID as well as the predicting factors and symptom features of long COVID in the primary care population and also including those who had not received a diagnosis or test confirmation for their symptoms. Due to the limited testing during the initial wave of the pandemic, most COVID-19 cases would have gone undetected and unrecorded in EHRs. This study, therefore, provides a more comprehensive picture of long COVID, which includes the previously undetected group of long COVID patients.

A previous investigation on persisting symptoms of COVID-19, which included untested patients, was conducted by Goertz et al.²⁴ However, in contrast to our study, participants in the study were recruited from online groups of patients with persisting symptoms. The current study further supplements our knowledge by including a broader population of primary care patients utilizing Platform C19's unique linkage to a quality improvement program and access to EHR data supplemented with patient-provided information. This provides the unique capability to answer research questions that are not possible with EHR data or questionnaires alone. The utility of patient-reported outcome measures had been previously utilized to demonstrate the social and physical impact of persisting COVID-19 symptoms.¹⁰

Limitations

Similar to other questionnaire studies, this study is prone to non-response bias. Patients who respond may not be fully representative of the general population. Due to the electronic nature of the questionnaire, it is possible that the elderly and frail were less likely to respond. The Black, Asian, and ethnic minorities are also currently underrepresented among our respondents.

The retrospective nature of this study may have introduced a recall bias. Patients whose symptoms lasted longer may perceive themselves as having more symptoms. Similarly, patients with more symptoms may have reported their symptoms to have lasted longer. This might have inflated the differences in symptom features between long COVID and COVID-19 of shorter duration. Lastly, symptom duration depends on patient recall as it is impossible to retrospectively confirm the length of COVID-19 related symptoms. We are also unable to ascertain whether the reported symptoms were present only during the acute phases of the infection or if they persisted as long-term symptoms.

Conclusion

This study provides novel insight on the burden and predictors of long COVID in a broad primary care population, which includes the population of previously undetected COVID-19 patients who were not clinically diagnosed or test-confirmed. These insights will be valuable for the identification of care pathways and interventions to support this patient group.

Long COVID was observed in almost 10% of COVID-19 cases and was mostly from the group of patients who did not receive test-confirmed COVID-19. Compared to test-confirmed COVID-19 patients, patients who were diagnosed by a clinician but not test-confirmed were significantly more likely to have long COVID. In contrast, patients who were self-diagnosed without clinician diagnosis nor test confirmation were significantly less likely. Older age, female sex, presence of frailty, visit to an A&E department, and hospital admission for COVID-19 symptoms or infection significantly predicted long COVID. Symptoms of aches and pain, appetite loss, confusion and disorientation, diarrhea, and persistent dry cough were indicative of cases that persisted into long COVID.

Further analyses using data stored within Platform C19 will look at the correlation between the risk of developing long COVID with the number of symptoms as a metric for disease severity, patient behavior, and psychological comorbidities. The multi-organ system symptomatology of COVID-19 is similar to that observed in functional disorders such as irritable bowel syndrome, fibromyalgia, and chronic fatigue syndrome.³¹ The potential co-morbid link between long COVID and functional disorders is a potential area for future analysis using Platform C19.

Data collection under the OPC COVID-19 QI program is ongoing. Future studies will investigate the long-term clinical outcome of long COVID and its effect on the well-being of patients. With the initiation of the vaccination program in the UK, questions on vaccination status will be included in future versions of the questionnaire. Future research using data stored in Platform C19 will analyze the protective effects of vaccinations on the development of long COVID.

Data Sharing Statement

Access to the de-identified patient data used in this study may be requested via the OPCRD website (<https://opcrd.co.uk/platform-c19-new-2/>) or via the inquiries email info@opcrd.co.uk.

Ethics

The OPCRDC has received NHS Research Ethics Committee (REC) approval to provide anonymized data for scientific and medical research since 2010, with its most recent approval in 2015 (NHS HRA REC ref: 15/EM/0150). OPCRDC is governed by the Anonymized Data Ethics and Protocols Transparency (ADEPT) committee (ADEPT1720).

OPC quality improvement programs are provided in compliance with the GDPR/Data Protection Act 2018 (Data Protection Register Ref: ZA197058) and the NHS Data Security and Protection Toolkit (Ref: 8HR85). The protocol of Platform C19 was approved by the ADEPT committee on 18th November 2020 (ADEPT1720). The study was approved on 18th November 2020 by the Anonymized Data Ethics and Protocols Transparency Committee (Ref: PROTOCOL2286; ADEPT ref: ADEPT1720).

Acknowledgments

Other members of the Platform C19 steering committee consist of Sir Prof Lewis Ritchie, OBE from the University of Aberdeen, Steve Davis from Interface Clinical Services, Dr Samantha Walker from Asthma UK & British Lung Foundation, Prof Stephen Holgate from the University of Southampton and the Medical Research Council, and Sue Beacroft from Open Vie/Harvey Walsh and OPCRDC-NEXUS.

We wish to acknowledge our GP Partner Practices: Alnwick Medical Group, Guide Post Medical Group, Priors Medical Group, Highgate Medical Centre, Clifton Medical Centre, Bridge Street Surgery, Park and St Francis Surgery, Brockworth Surgery, Whitmoor Medical Centre, Tong Medical Practice, Eric Moore Partnership, Bridgewater Surgeries, The Village Surgery, and Queen Square Medical Practice.

Author Contributions

All authors contributed to data analysis, drafting or revising the article, gave final approval of the version to be published, agreed to the submitted journal, and agree to be accountable for all aspects of the work. The overall conduct of this study was supervised by David Price.

Funding

The design, conduct, and writing of this study are funded by Optimum Patient Care UK and the Observational and Pragmatic Research Institute Singapore.

Disclosure

Rupert Jones reports grants, personal fees, and non-financial support from AstraZeneca and OPRI, personal fees and non-financial support from Boehringer Ingelheim, grants, personal fees, and non-financial support from GSK, grants and non-financial support from Novartis, non-financial support from Nutricia, and personal fees from Pfizer outside the submitted work.

Andrew Davis was an employee of Observational and Pragmatic Research Institute at the time this study was conducted.

Brooklyn Stanley and Anu Kemppinen are employees of Observational and Pragmatic Research Institute, which conducted this study and conducted paid research in respiratory disease on behalf of the following other organizations in the past 5 years: Aerocrine, AKL Research and Development Ltd, Almirall, AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Mylan, Mundipharma, Napp, Novartis, Orion, Takeda, Teva, Zentiva (a Sanofi company).

Dermod Ryan has (in the last 3 years) lectured on behalf of, received sponsorship from, or acted as a paid advisor to Mylan, AZ, Chiesi, Novartis, GSK, Boehringer Ingelheim and Regeneron.

David J Jackson has received advisory board and speaker fees from AstraZeneca, GSK, BI, Teva, Napp, Chiesi, Novartis and research grant funding from AstraZeneca.

David MG Halpin has received sponsorship to attend international meetings, and honoraria for lecturing, attending advisory boards and preparing educational materials from AstraZeneca, Boehringer Ingelheim, Chiesi, GSK, Novartis and Pfizer.

Jennifer K Quint reports grants from MRC, grants from The Health Foundation, grants and personal fees from AZ, grants from Bayer, grants and personal fees from Chiesi, grants and personal fees from GSK, grants and personal fees from BI, outside the submitted work.

Kamlesh Khunti has been a consultant and speaker for Novartis, Novo Nordisk, Sanofi-Aventis, Lilly, and Merck Sharp & Dohme; has received grants in support of investigator-initiated trials from Novartis, Novo Nordisk, Sanofi-Aventis, Lilly, Merck Sharp & Dohme, Pfizer, and Boehringer Ingelheim; has served on advisory boards for Novo Nordisk, SanofiAventis, Lilly, and Merck Sharp & Dohme; and is supported by the UK National Institute for Health Research (NIHR) Applied Research Collaboration East Midlands and the NIHR Leicester Biomedical Research Centre.

Liam G Heaney is an Academic Lead for the UK MRC Consortium for Stratified Medicine in Severe Asthma – Industrial Pharma partners Amgen, AstraZeneca, MedImmune, Janssen, Novartis, Roche/Genentech, GlaxoSmithKline, and Boehringer Ingelheim; prior project grant funding from MedImmune, Novartis UK, Roche/Genentech, and GlaxoSmithKline; has taken part in Advisory Boards/Lectures supported by Novartis, Roche/Evelo Biosciences, Genentech, GlaxoSmithKline, Teva, Theravance and Vectura; has travel funding support to international respiratory meetings (AstraZeneca, Chiesi, Novartis, Boehringer Ingelheim, Teva, and GlaxoSmithKline) and has taken part in asthma clinical trials (GlaxoSmithKline, Schering Plough, Synairgen, Novartis, and Roche/Genentech) for which his institution was remunerated.

Salman Siddiqui reports advisory board/advisory services and speaker fees from AstraZeneca, GlaxoSmithKline, Chiesi, Boehringer Ingelheim, Novartis, Mundipharma, ERT medical, Owlstone Medical and received grants from UKRI PHOSP-COVID consortium.

Ian Pavord reports grants from NIHR and personal fees from Aerocrine, Almirall, Boehringer Ingelheim, Chiesi, Circassia, Genentech, GlaxoSmithKline, Knopp, Novartis, Regeneron, Sanofi and Teva, outside the submitted work, and is a member of the GOLD Science Committee. Michael Hyland has received personal fees from Novartis and GSK.

Stephen T Holgate is a co-founder of Synairgen Research, owns shares in the company, and has received personal fees as a Non-Executive Director and Consultant of Synairgen Research outside of the submitted work.

Francis Appiagyei, Emma-Jane Roberts, Megan Preston, Marije van Melle and Victoria Carter are employees of Optimum Patient Care who funded this study.

David Price has board membership with AstraZeneca, Boehringer Ingelheim, Chiesi, Mylan, Novartis, Regeneron Pharmaceuticals, Sanofi Genzyme, Thermofisher; consultancy agreements with Airway Vista Secretariat, AstraZeneca, Boehringer Ingelheim, Chiesi, EPG Communication Holdings Ltd, FIECON Ltd, Fieldwork International, GlaxoSmithKline, Mylan, Mundipharma, Novartis, OM Pharma SA, PeerVoice, Phadia AB, Spirosure Inc, Strategic North Limited, Synapse Research Management Partners S.L., Talos Health Solutions, Theravance and WebMD Global LLC; grants and unrestricted funding for investigator-initiated studies (conducted through Observational and Pragmatic Research Institute Pte Ltd) from AstraZeneca, Boehringer

Ingelheim, Chiesi, Mylan, Novartis, Regeneron Pharmaceuticals, Respiratory Effectiveness Group, Sanofi Genzyme, Theravance and UK National Health Service; payment for lectures/speaking engagements from AstraZeneca, Boehringer Ingelheim, Chiesi, Cipla, GlaxoSmithKline, Kyorin, Mylan, Mundipharma, Novartis, Regeneron Pharmaceuticals and Sanofi Genzyme; payment for travel/accommodation/meeting expenses from AstraZeneca, Boehringer Ingelheim, Mundipharma, Mylan, Novartis, Thermofisher; stock/stock options from AKL Research and Development Ltd, which produces phytopharmaceuticals; owns 74% of the social enterprise Optimum Patient Care Ltd (Australia and UK) and 92.61% of Observational and Pragmatic Research Institute Pte Ltd (Singapore); 5% shareholding in Timestamp, which develops adherence monitoring technology; is peer reviewer for grant committees of the UK Efficacy and Mechanism Evaluation programme, and Health Technology Assessment; and was an expert witness for GlaxoSmithKline.

Michael Hyland reports personal fees from GSK, outside the submitted work.

Pam Young and Tony Megaw are employees of Wellbeing Software.

The remaining authors report no conflicts of interest.

References

- Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med.* 2021;27:601.
- Venkatesan P. NICE guideline on long COVID. *Lancet Respir Med.* 2021;9(2):129. doi:10.1016/S2213-2600(21)00031-X
- National institute for health and care excellence: clinical guidelines. In: *COVID-19 Rapid Guideline: Managing the Long-Term Effects of COVID-19*. London: National Institute for Health and Care Excellence (UK);2020.
- Perego E, Callard F, Stras L, Melville-Jóhannesson B, Pope R, Alwan N. Why the patient-made term 'long covid' is needed [version 1; peer review: 1 approved with reservations, 1 not approved]. *Wellcome Open Res.* 2020;5(224).
- Greenhalgh T, Knight M, A'Court C, Buxton M, Husain L. Management of post-acute covid-19 in primary care. *BMJ.* 2020;370:m3026. doi:10.1136/bmj.m3026
- Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *MedRxiv.* 2020.
- Sykes DL, Holdsworth L, Jawad N, Gunasekera P, Morice AH, Crooks MG. Post-COVID-19 symptom burden: what is long-COVID and how should we manage it? *Lung.* 2021;192:113.
- Townsend L, Dyer AH, Jones K, et al. Persistent fatigue following SARS-CoV-2 infection is common and independent of severity of initial infection. *PLoS One.* 2020;15(11):e0240784. doi:10.1371/journal.pone.0240784
- Carfi A, Bernabei R, Landi F. Gemelli against C-P-ACSG. Persistent symptoms in patients after acute COVID-19. *JAMA.* 2020;324(6):603–605. doi:10.1001/jama.2020.12603

10. Ganesh R, Ghosh AK, Nyman MA, et al. PROMIS scales for assessment of persistent post-COVID Symptoms: a cross sectional study. *J Prim Care Community Health*. 2021;12:21501327211030413. doi:10.1177/21501327211030413
11. Tenforde MW, Kim SS, Lindsell CJ, et al. Symptom duration and risk factors for delayed return to usual health among outpatients with COVID-19 in a multistate health care systems network -United States, March-June 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(30):993–998. doi:10.15585/mmwr.mm6930e1
12. Halpin SJ, McIvor C, Whyatt G, et al. Postdischarge symptoms and rehabilitation needs in survivors of COVID-19 infection: a cross-sectional evaluation. *J Med Virol*. 2021;93(2):1013–1022. doi:10.1002/jmv.26368
13. Maestre-Muñiz MM, Arias Á, Mata-Vázquez E, et al. Long-term outcomes of patients with coronavirus disease 2019 at one year after hospital discharge. *J Clin Med*. 2021;10(13):2945. doi:10.3390/jcm10132945
14. The prevalence of long COVID symptoms and COVID-19 complications [Website]. Office for national statistics. Available from: <https://www.ons.gov.uk/news/statementsandletters/theprevalenceoflongcovid-symptomsandcovid19complications>. Accessed February 2, 2021.
15. Jacobson KB, Rao M, Bonilla H, et al. Patients with uncomplicated COVID-19 have long-term persistent symptoms and functional impairment similar to patients with severe COVID-19: a cautionary tale during a global pandemic. *Clin Infect Dis*. 2021. doi:10.1093/cid/ciab103
16. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. *Nat Med*. 2021;27(4):626–631. doi:10.1038/s41591-021-01292-y
17. Böhning D, Rocchetti I, Maruotti A, Holling H. Estimating the undetected infections in the Covid-19 outbreak by harnessing capture–recapture methods. *Int J Infect Dis*. 2020;97:197–201. doi:10.1016/j.ijid.2020.06.009
18. Rocchetti I, Böhning D, Holling H, Maruotti A. Estimating the size of undetected cases of the COVID-19 outbreak in Europe: an upper bound estimator. *J Epidemiol Methods*. 2020;9:s1.
19. Arnold D, Hamilton F, Milne A, et al. Patient outcomes after hospitalisation with COVID-19 and implications for follow-up; results from a prospective UK cohort. *Thorax*. 2021;76(4):399. doi:10.1136/thoraxjnl-2020-216086
20. Simani L, Ramezani M, Darazam IA, et al. Prevalence and correlates of chronic fatigue syndrome and post-traumatic stress disorder after the outbreak of the COVID-19. *J Neurovirol*. 2021;27(1):154–159. doi:10.1007/s13365-021-00949-1
21. Chopra V, Flanders SA, O'Malley M, Malani AN, Prescott HC. Sixty-day outcomes among patients hospitalized with COVID-19. *Ann Intern Med*. 2020;174:576.
22. Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet (London, England)*. 2021;397(10270):220–232. doi:10.1016/S0140-6736(20)32656-8
23. Evans RA, McAuley H, Harrison EM, et al. Physical, cognitive and mental health impacts of COVID-19 following hospitalisation – a multi-centre prospective cohort study. *MedRxiv*. 2021.
24. Goertz YMJ, Van Herck M, Delbressine JM, et al. Persistent symptoms 3 months after a SARS-CoV-2 infection: the post-COVID-19 syndrome? *ERJ Open Res*. 2020;6(4):00542–2020. doi:10.1183/23120541.00542-2020
25. National data opt-out [website]. NHS digital; 2021. Available from: <https://digital.nhs.uk/services/national-data-opt-out/compliance-with-the-national-data-opt-out/check-for-national-data-opt-outs-service>. Accessed July 16, 2021.
26. Peckham H, de Groot NM, Raine C, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ICU admission. *Nat Commun*. 2020;11(1):6317. doi:10.1038/s41467-020-19741-6
27. Ladwig KH, Marten-Mittag B, Formanek B, Dammann G. Gender differences of symptom reporting and medical health care utilization in the German population. *Eur J Epidemiol*. 2000;16(6):511–518. doi:10.1023/A:1007629920752
28. Hewitt J, Carter B, Vilches-Moraga A, et al. The effect of frailty on survival in patients with COVID-19 (COPE): a multicentre, European, observational cohort study. *Lancet Public Health*. 2020;5(8):e444–e451. doi:10.1016/S2468-2667(20)30146-8
29. COVID-19 rapid guideline: critical care in adults. National institute for clinical excellence; 2021. Available from: <https://www.nice.org.uk/guidance/ng159/resources/covid19-rapid-guideline-critical-care-in-adults-pdf-66141848681413>. Accessed July 16, 2021.
30. Chopra N, Chowdhury M, Singh AK, et al. Clinical predictors of long COVID-19 and phenotypes of mild COVID-19 at a tertiary care centre in India. *Drug Discov Ther*. 2021;15(3):156–161. doi:10.5582/ddt.2021.01014
31. Hyland ME, Bacon AM, Lanario JW, Davies AF. Symptom frequency and development of a generic functional disorder symptom scale suitable for use in studies of patients with irritable bowel syndrome, fibromyalgia syndrome or chronic fatigue syndrome. *Chronic Dis Transl Med*. 2019;5(2):129–138. doi:10.1016/j.cdtm.2019.05.003

Pragmatic and Observational Research

Dovepress

Publish your work in this journal

Pragmatic and Observational Research is an international, peer-reviewed, open access journal that publishes data from studies designed to reflect more closely medical interventions in real-world clinical practice compared with classical randomized controlled

trials (RCTs). The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/pragmatic-and-observational-research-journal>