Clinical characteristics with inflammation profiling of Long-COVID and association with one-year recovery following hospitalisation in the UK: a prospective observational study

On behalf of the PHOSP-COVID Collaborative Group†

†The full writing group is listed at the end of the manuscript and a complete list of members of the PHOSP-COVID Collaborative Group is provided in the supplement

Correspondence to: Prof Christopher E Brightling

Institute for Lung Health,

Department of Respiratory Sciences,

NIHR Leicester BRC,

University of Leicester, Leicester, United Kingdom

phosp@leicester.ac.uk

Funding Jointly funded by UK Research and Innovation and National Institute of Health Research (grant references: MR/V027859/1 and COV0319).

Ethics Approval Ethics Ref: 20/YH/0225

Summary

Background

There are currently no effective pharmacological or non-pharmacological interventions for Long-COVID. We aimed to describe recovery one year after hospital discharge for COVID-19 and identify factors associated with patient-perceived recovery. To identify potential therapeutic targets, we focussed on previously described four recovery clusters five months after hospital discharge, their underlying inflammatory profiles and relationship with clinical outcomes at one year.

Methods

PHOSP-COVID is a prospective longitudinal cohort study, recruiting adults hospitalised with COVID-19 across the UK. Recovery was assessed using patient reported outcomes measures (PROMs), physical performance, and organ function at five-months and one-year after hospital discharge. Hierarchical logistic regression modelling was performed for patient-perceived recovery at one-year. Cluster analysis was performed using clustering large applications (CLARA) k-medoids approach using clinical outcomes at five-months. Inflammatory protein profiling from plasma at the five-month visit was performed.

Findings

2320 participants have been assessed at five months after discharge and 807 participants have completed both five-month and one-year visits. Of these, 35·6% were female, mean age 58·7 (SD 12·5) years, and 27·8% received invasive mechanical ventilation (IMV). The proportion of patients reporting full recovery was unchanged between five months 501/1965 (25·5%) and one year 232/804 (28·9%). Factors associated with being less likely to report full recovery at one year were: female sex OR 0·68 (95% CI 0·46-0·99), obesity OR 0·50 (95%CI 0·34-0·74) and IMV OR 0·42 (95%CI 0·23-0·76).

Cluster analysis (n=1636) corroborated the previously reported four clusters: ‘very severe’, ‘severe’, ‘moderate/cognitive’, ‘mild’ relating to the severity of physical, mental health and cognitive impairments at five months in a larger sample. There was elevation of inflammatory mediators of tissue damage and repair in both the 'very severe' and the 'moderate/cognitive' clusters compared to the 'mild' cluster including interleukin-6 which was elevated in both comparisons (analysis with total n=626 participants). Overall, there was a substantial deficit in median (IQR) EQ5D-5L utility index from pre-COVID (retrospective assessment) 0·88 (0·74-1·00), five months 0·74 (0·60-0·88) to one year: 0·74 (0·59-0·88), with minimal improvements across all outcome measures at one-year after discharge in the whole cohort and within each of the four clusters.

Interpretation

The sequelae of a hospital admission with COVID-19 remain substantial one year after discharge across a range of health domains with the minority in our cohort feeling fully recovered. Patient perceived health-related quality of life remains reduced at one year compared to pre-hospital admission. Systematic inflammation and obesity are potential treatable traits that warrant further investigation in clinical trials.

Funding UKRI & NIHR

Research in Context

Evidence before this study

We systematically searched PubMed and Embase databases for large studies reporting one-year follow-up data for hospitalised COVID-19 patients published between January 1, 2021 and November 7, 2021, without language restrictions. Search terms related to COVID-19, hospitalisation and long-term follow-up were used. A large prospective cohort study from Wuhan, China (n = 1276) showed that 49% of patients reported at least one persistent symptom during a follow-up clinic visit at 12 months post COVID-19; no significant improvement in exercise capacity was observed between six- and 12-month visits. Another two large cohort studies in China (n = 2433) and Spain (n = 1950) with one-year follow-up data from telephone interviews showed that 45% and 81% of patients reported at least one residual COVID-19 symptom, respectively. However, no previous studies have compared the trajectories of COVID-19 recovery in patients classified by different clinical phenotypes, and there are no large studies investigating the relationship between systemic inflammation and ongoing health impairments post COVID-19.

Added value of this study

In a diverse population of adults post-hospital admission with COVID-19, our large UK prospective multi-centre study reports several novel findings: the minority felt fully recovered at one year with minimal recovery from five months across any health domain; female sex and obesity are associated with being less likely to feel fully recovered at one year; several inflammatory mediators were increased in individuals with the most severe physical, mental health, and cognitive impairments compared to individuals with milder ongoing impairments.

Implications of all the available evidence

Both pharmacological and non-pharmacological interventions are urgently needed to improve the ongoing burden following hospitalisation for COVID-19 both for individuals and

92 healthcare systems; our findings support the use of a precision medicine approach with
93 potential treatable traits of systemic inflammation and obesity.

Introduction

To date (December, 2021), there have been over 260 million cases of SARS-CoV-2 reported worldwide,[1] 10 million cases in the UK[2] and over half a million patients in the UK admitted to hospital for COVID-19. We, and others, have shown that this sizeable population are at high risk of persisting health impairments six months after discharge associated with reduced physical function and health-related quality of life.[3-4] It is essential to understand the longer-term trajectory of recovery to identify ongoing healthcare needs, and the required response by healthcare systems and policy makers for this already large and ever-increasing population.

Much remains unknown about the longer-term sequelae of COVID-19. In the largest cohort study to date from Wuhan, China, nearly half of patients experienced persistent symptoms 12 months post-discharge from hospital for COVID-19.[5] Between six- and 12-months post-discharge, there was no change in six-minute walk distance, but some improvement in the results of pulmonary imaging.[5]

The mechanisms underlying long-term persistence of symptoms are currently unknown. A potential hypothesis is that the hyperinflammation associated with acute COVID-19 leads to a persistent inflammatory state post COVID-19, associated with dysregulated immunity and multiorgan dysfunction. Although multiple studies have highlighted elevated inflammatory markers, including IL-6, associated with severity of acute illness,[6,7] there are no large studies investigating the relationship between systemic inflammation and ongoing health impairments after COVID-19.

Currently, there are no effective treatments for Long-COVID/ post-COVID-19 condition. Long-COVID is defined by National Institute for Health and Care Excellence (NICE) as on-

119 going symptoms beyond four to 12 weeks after COVID-19 and post-COVID-19 condition by
120 the World Health Organisation as occuring "in individuals with a history of probable or
121 confirmed SARS CoV-2 infection, usually three months from the onset of COVID-19 with
122 symptoms and that last for at least two months and cannot be explained by an alternative
123 diagnosis"*)*.[8,9] Improved characterisation of this population with an emphasis on elucidating
124 underlying mechanisms is needed to identify potential therapeutic targets. We previously
125 described four clinical recovery clusters of patients: 'very severe', 'severe',
126 'moderate/cognitive', and 'mild' defined by severity of ongoing physical, mental health and
127 cognitive impairment five months after a hospital admission with COVID-19. We sought to
128 answer the following questions using the ongoing PHOSP-COVID longitudinal study cohort:
129 firstly, what proportion of patients discharged from hospital with COVID-19 felt fully
130 recovered one-year later and what are the characteristics associated with non-recovery.
131 Secondly, are there inflammatory mediators associated with severity of ongoing health
132 impairments and therefore potential therapeutic targets? Thirdly, are there differences in the
133 trajectory of recovery at one year across different health domains and between our previously
134 described clusters?

Methods

Study design and participants

The PHOSP-COVID cohort recruitment strategy including eligibility criteria has been described previously.[3] In brief, we recruited patients aged 18 years and older who were discharged from over 50 National Health Service (NHS) hospitals across the four UK nations following admission to a medical assessment unit or ward for confirmed or clinician-diagnosed COVID-19 before March 31, 2021. The current analysis involves participants who consented to attend two additional in-person research visits (Tier 2 – see protocol and Table S1) within one-year after discharge alongside routine clinical care.

Written informed consent was obtained from all study participants. The study was approved by the Leeds West Research Ethics Committee (20/YH/0225) and is registered on the ISRCTN Registry (ISRCTN10980107).

Procedures

Participants were invited to attend research visits at two to seven months post-discharge ('five-month visit') and at 10 to 14 months ('one-year visit'). Participants were also able to attend a one-year only visit if they were outside the time period for a five-month visit at the time of consent and were discharged before November 30, 2020. The core set of data variables collected at each visit and included in this study are listed in the Supplementary materials, Table S2. These included baseline demographics, information about disease severity and treatment during their hospital admission, as well as symptoms using a bespoke study-specific questionnaire and other patient reported outcome measures (PROMs) for anxiety (Generalised Anxiety Disorder [GAD-7]), depression (Patient Health Questionnaire [PHQ-9]), Post-traumatic stress disorder CheckList for Diagnostic and Statistical Manual of Mental

Disorders (DSM-5) [PCL-5], fatigue (Functional Assessment of Chronic Illness Therapy – Fatigue [FACIT-Fatigue]), breathlessness (Dyspnoea-12), and health-related quality of life (EQ5D-5L), physical performance measures including the short physical performance battery (SPPB) and the incremental shuttle walk test (ISWT), cognitive impairment using the Montreal Cognitive Assessment (MoCA), pulmonary function tests and blood test results reflecting multi-organ function and systemic inflammation obtained at their clinical and research visits (see Supplementary methods). Patients were also asked to complete the EQ5D-5L, Washington Group Short Set Functioning (WG-SS) and Visual analogue scale for Breathlessness and Fatigue retrospectively to assess their perceived pre-COVID health (Supplementary Table S2). Plasma samples obtained at the five-month visit were analysed using the Olink Explore 384 Inflammation panel. Sample processing and assay details are provided in the Supplementary methods.

The primary outcome for this analysis was patient-perceived recovery assessed using a study-specific questionnaire and the question 'Do you feel fully recovered?' and participants could answer 'Yes', 'No', or 'Not sure'. Other secondary outcomes included symptoms since their COVID-19 hospital admission collected on the bespoke study-specific questionnaire, validated patient reported outcome questionnaires, and physiological measures (including physical performance and spirometry). See Supplementary Tables S1 and S2 for details of all variables included in this study.

Statistical analysis

Continuous variables were presented as median and interquartile range (IQR) or mean and standard deviation (SD). Binary and categorical variables were presented as counts and

percentages (by row or by column as indicated in table legends). Participants were stratified by patient-perceived recovery: Yes (recovered), Not sure or No (not recovered).

Missing data were reported within each variable and per category. Within visit, a chi-squared test was used to identify differences in proportions across multiple categories. For normally distributed and non-normally distributed continuous data, analysis of variance (ANOVA F-test) and Kruskal Wallis tests respectively, were used to test differences across categories. For paired data between the five-month and one-year visit, a McNemar's chi-squared test with continuity correction and a McNemar's chi-squared test were used for binary variables and variables with > 2 levels, respectively. For normally distributed and non-normally distributed continuous data, a paired t-test and a Wilcoxon signed-rank test were used, respectively. As previously described,[3] univariable and hierarchical multivariable logistic regression models (admission hospital included as random effect) were used to explore risk factors associated with patient-perceived recovery. Missing data was addressed using multiple imputation (10 datasets, 10 iterations, final models combined using Rubin's Rules), with the outcome used in imputation models, but not itself imputed.

In order to evaluate any potential bias due to patients not yet attending their one-year visit at the time of analysis (October 6, 2021), we compared characteristics and patient-perceived recovery between those who attended a one-year visit versus those who had had not yet attended but were discharged from hospital during the same range of dates. Multiple imputation was used to complete missing outcomes for participants who had not yet attended their one-year follow-up. The imputation model used age, sex, ethnicity, Index of Multiple Deprivation, and WHO clinical progression scaleand all comorbidity variables. Ten datasets with 10 iterations were created, and combined using Rubin's rules.

In this larger cohort, we repeated our previous unsupervised cluster analysis[3] of patient recovery measured using symptom questionnaires (PROMs), physical performance and cognitive assessment data (Questionnaires: Dyspnoea-12, FACIT-Fatigue, GAD-7, PHQ-9, PCL-5, SPPB and MoCA as continuous variables) from the five-month visit (discharge dates Feb 2020 – March 2021) using the clustering large applications (CLARA) k-medoids approach.[10] Scores were centred, normalised and transformed so higher burden of disease represented higher values. A Euclidean distance metric was used, and the optimal number of clusters chosen using a silhouette plot. Cluster membership was determined for each individual using five-month visit data. Characteristics at one-year and change in characteristics between five-months and 12-months, are presented as cluster-stratified tables. All tests were two-tailed and p values <0·05 were considered statistically significant. We did not adjust for multiple testing.

Plasma protein levels were compared between clusters using the mildest recovery cluster as baseline and using multinomial regression with age, body mass index (BMI) and number of comorbidities as covariates (see Supplementary Methods). Significance was defined as p <0·1 after False Discovery Rate adjustment for multiple testing.

We used R (version 3·6·3) with the *finalfit*, *tidyverse*, *mice*, *cluster*, *ggplot2*, *ggalluvial*, *radiant*, *dabestr* and *recipes* packages.

Results

At the time of analysis October 6, 2021, 2320 participants (discharged from hospital March 7, 2020 to April 18, 2021) had attended a five-month visit (median [IQR] 5 [4-6] months post-discharge), 924 participants (discharged February 28, 2020 to November 28, 2020) had returned for a one-year visit (13 [12-13]) months post-discharge) and 807 participants had attended both visits (Figure 1). The individual and hospital admission characteristics including severity of acute illness between those with a five-month visit, one-year visit, and those who attended both visits were similar (Table 1) except for acute treatment with corticosteroids.

Proportion of patients feeling fully recovered and characteristics associated with patient-perceived recovery at one year

Figure 1 details the number of participants that had both attended a research visit and answered whether or not they felt fully recovered: n=1965 at five months and n=804 at one year . At five-months post-discharge, 25·5% (501/1965) of patients felt fully recovered with 19·6% (385/1965) feeling not sure and 54·9 (1079/1965) definitely not recovered (Figure 2a, Table S3a).At one year 28·9% (232/804) of patients felt fully recovered, 22·4% (180/804) were not sure and 48·8 (392/804) were not recovered (Figure 2a Table S3b). Similar proportions were observed in those with paired data (Table S3b). The individual responses were also similar between five-months and one-year shown in an alluvial plot Figure S1.

In multivariable analysis, female sex (OR 0·68 (0·46-0·99)), Body Mass Index (BMI) 30 kg/m^2 or greater (OR 0·50 (0·34-0·74)) and receiving invasive mechanical ventilation (IMV) World Health Organisation [WHO] category 7-9 (OR 0·42 (0·23-0·76)) were all independent factors associated with being less likely to recover at one-year (Figure 2b and Table S). There was no effect of receiving systemic corticosteroids (OR 1·05 (0·66-1·65)) during the acute admission on patient-perceived recovery at one-year for the whole cohort (Figure 2b, Table S4). There was also no effect of time from discharge to the research visit (OR 1·00 (1·00-1·01)).

751 participants discharged between February 28, 2020 and November 28, 2020 had not yet returned for a one-year visit but had similar characteristics and five-month recovery status to the 924 participants who had attended (Tables S5). The proportion of recovered patients was similar after imputation for outcome 499/1675 (29·8%).

Inflammatory mediators associated with four recovery clusters

For the current five-month dataset, the previously identified four clusters[3] were confirmed using participants with complete data for the cluster analysis n=1636 (Figure 1). The distribution of the four clusters were 'very severe' physical and mental health impairment (n=319), 'severe' physical and mental health impairment (n=493), 'moderate/cognitive' physical health impairment with cognitive impairment (n=179) and 'mild' mental and physical health impairment (n=645) (Table S6, Figure S2). 86·7% (664/766) of individuals included in the previous study[3] were re-assigned to the same recovery cluster as before; the cognitive cluster had the most assignment alterations (60/127). Characteristics of individuals in each recovery cluster are shown in Table S7. The very severe cluster had a higher proportion of female sex (165/306, 53·9%) and obesity (204/288, 70·8%) compared to the 'mild' cluster (177/624, 28·4%) and (288/568, 50·2%), respectively.

After quality control, plasma proteome data for 296 protein features and complete clinical data for cluster assignment were available at five-months for 626 participants: 'very severe' cluster n=111 participants, 'severe' cluster n=173 participants, 'moderate/cognitive' cluster n=73 participants and 'mild' cluster n=269 participants. Age, BMI, and two or more co-morbidities were associated with cluster membership whereas receiving IMV during the acute illness was not (analysis in participants with plasma proteome data and a cluster assignment) (Table S8). After adjustment for age, BMI, and co-morbidity count, 13 proteins were significantly elevated

in participants in the 'very severe' recovery cluster compared to those in the 'mild' cluster (Table S9, Figure 3). These were Trefoil Factor 2 (TFF2), Transforming Growth Factor Alpha (TGFA), Lysosomal Associated Membrane Protein 3 (LAMP3), CD83 molecule (CD83), Galectin-9 (LGALS9), Plasminogen Activator, Urokinase Receptor (PLAUR), Interleukin 6 (IL6), Erythropoietin (EPO), Fms Related Receptor Tyrosine Kinase 3 Ligand (FLT3LG), Agrin (AGRN), Secretoglobin Family 3A Member 2 (SCGB3A2), Follistatin (FST) and C-Type Lectin Domain Family 4 Member D (CLEC4D) (Figure S3). In addition, IL6 and CD70 molecule (CD70) were significantly increased in the 'moderate/cognitive' recovery cluster compared to the 'mild' cluster.

One year recovery across symptoms, health domains and the four recovery clusters

The top 10 most common persistent symptoms at one-year post-discharge were fatigue (463/770, 60·1%), aching muscles (442/809, 54·6%), physical slowing down (429/811, 52·9%), poor sleep (402/769, 52·3%), breathlessness (395/769, 51·4%), joint pain or swelling (382/803, 47·6%), slowing down in thinking (377/808, 46·7%), pain (359/770, 46·6%), short term memory loss (360/808, 44·6%) and limb weakness (341/813, 41·9%) (Table S10). Overall, these were essentially unchanged in prevalence from five months to one-year, with small reductions in rates of limb weakness (47·6% at five-months vs. 41·7% at one-year, p=0·010), paraesthesia (40·6% vs. 35·2%, p=0·014) and balance problems (34·9% vs. 30·0, p=0·008). There was either no or minimal improvement in PROMS, physical function, cognitive impairment, or organ function at one-year compared to five-months post-discharge (paired data Table 2 and presented stratified by patient-perceived recovery in Tables S3a and b). At one year, 147/684 (21·5%) and 169/680 (24·9%) participants had clinically relevant symptoms of anxiety and/or depression, respectively, 68/680 (10·0%) had symptoms compatible with post-traumatic stress disorder, and 55/623 (8·8%) had significant cognitive

impairment (Table 2). Measures of symptoms and physical function were significantly different across participants feeling fully recovered, not sure, or not fully recovered at five months and one-year but cognitive impairment and measures of organ function (except for Forced Vital Capacity [FVC]) were not (Table 2). Health-related quality of life was significantly different across participants reporting being fully recovered, not sure or not recovered at both five months, and one year (Figure 2d and Tables S3a and b).

In addition to higher proportions of female sex and obesity (Table S7), at one year the more severe clusters were associated with lower proportion of feeling fully recovered, 11/272 (4·0%) vs. 90/179 (50·3%) (Figure 2c), reduced exercise capacity ISWT 44·4% predicted vs. 72·4% predicted, higher number of symptoms 20 vs 4, and higher proportion of elevated CRP level >5mg/L 38·4% vs. 14·5% compared to the mild cluster (Table 3, Figure 4a). A comparison of health outcomes across the four clusters between the five-month and one-year time-points (n=602) shows there was minimal change across the two time-points for the four clusters (Table 3). In the 'very severe' cluster, symptoms of anxiety, depression, breathlessness and fatigue significantly improved between five-months and one-year, but with minimal change in physical performance and no overall change in systemic inflammation measured by CRP levels. Cognitive impairment significantly improved at one-year in the 'moderate/cognitive' cluster and showed a non-significant trend towards improvement in the 'severe' cluster but was unchanged in the other clusters (Table 3). Compared with patient perceived pre-COVID-19 health, decrements were seen at five months and sustained at one-year across health-related quality of life (EQ5D-5L) Figure 4b, disability (WG-SS-SCo), and severity of breathlessness and fatigue experienced in the last 24 hrs (, Table S11, Figure S4, Figure S5).

Discussion

In a diverse UK population of adult survivors of COVID-19, we found the minority of participants felt fully recovered one year after hospital discharge with minimal improvement after their five-month assessment. The most common ongoing symptoms were fatigue, muscle pain, physically slowing down, poor sleep and breathlessness. The major risk factors for failure to recover at one-year were female sex, obesity and receiving IMV during the acute illness. There were substantial impairments in health-related quality of life (HRQoL) at five-months and one-year compared to retrospective self-reported pre-infection levels. Cluster analysis using the participants' five-month assessment corroborated four different clusters: 'very severe', 'severe', 'moderate/cognitive', and 'mild' based on the severity of physical, mental and cognitive impairments with similar characteristics to previously reported.[3] We confirmed that obesity, reduced exercise capacity, a higher number of symptoms and elevated serum CRP level were associated with the more severe clusters.[3] In the largest post-hospital cohort with systemic inflammatory profiling to date, inflammatory mediators consistent with persistent lung and systemic inflammation were elevated between both the 'very severe' and 'moderate/cognitive' clusters, and the 'mild' cluster. We therefore highlight traits to identify individuals at high risk of non-recovery and potential targetable pathways for interventions.

Comparing the systemic inflammatory profiling at five months after discharge between the 'very severe' and 'mild' cluster, the most elevated protein, TFF2, is a protein released with mucin from mucosal epithelium including lung and gastric mucosa. It has postulated roles in repair of damaged epithelium[11] and in combination with IFN-kappa reduced duration of infection in a small open-label randomised controlled trial of acute COVID-19.[11] In a previous study of patients during acute illness with COVID-19 using Olink proteomics, IL-6 was the most upregulated protein at day seven amongst patients who developed ARDS and

subsequently died.[6] Similarly, other proteins we identified such as LAMP3, Gal-9, CD83 are involved in T-cell macrophage and dendritic cell activation and were associated with increased morbidity and mortality during acute COVID-19 infection.[12-14] These changes suggest persistent mucosal epithelial abnormalities and inflammatory cell activation. Elevated serum levels of the C-terminal fragment of Agrin have been reported in older adults with sarcopenia, possibly related to break down of the neuromuscular junction.[15] The raised Agrin levels seen here may therefore contribute to the high prevalence of physical impairment. Interestingly, in the 'moderate/cognitive' cluster versus the 'mild' cluster IL6 and CD70 were elevated suggesting possible neuroinflammation contributing to the cognitive impairment as CD70 has been implicated in inflammation in the Central Nervous System (CNS)[16] via a role in differentiation of proinflammatory pathogenic lymphocytes. There were small improvements in cognition in the 'moderate/cognitive' cluster and a trend towards improvement in the 'very severe' cluster indicating that some of this deficit was not pre-existing and potentially modifiable; however, considerable deficit persisted at one year. The associations with the inflammatory mediators remained after adjusting for age, BMI and co-morbidities, and the proportion having received invasive mechanical ventilation was similar across the clusters; all factors known to be associated with systemic inflammation.[17] Taken together, the elevated mediators provide biological plausibility for the persistent severe-impairments seen in physical, mental health and cognitive impairment after COVID-19.

The limited recovery from five-months to one-year post-hospitalisation in our study across symptoms, mental health, exercise capacity, organ impairment and quality-of-life is striking. There are limited similar detailed prospective longitudinal studies for patients hospitalised with COVID-19 but in a larger cohort we confirm their findings of minimal recovery.[18-20] Although the large-scale study from Wuhan, China suggests a greater magnitude of recovery, new onset

symptoms persisted in half of the patients.[5] Notably, the Wuhan cohort included less severe acute illness with only 1% requiring IMV and 7% requiring high flow nasal oxygen and/or continuous positive airway pressure (CPAP), had fewer pre-existing co-morbidities, and a higher proportion of never-smokers. In non-COVID related ARDS survivors, little recovery in health-related quality of life is observed beyond six months, but larger improvements in walking distance[21,22] than we report post-COVID-19 in our cohort where over 70% did not receive IMV. In non-hospitalised patients after COVID-19 the proportion that develop Long-COVID appears lower.[23,24]

The responses for patient perceived recovery were discriminatory across all the PROMS and exercise measures providing additional validity for this outcome measure. We found female sex and obesity were major risk factors for not recovering at one year confirming existing results from smaller cohorts[25,] and non-hospitalised cohorts.[26-28] Female sex was similarly associated with worse recovery for fatigue, mental health and lung function at 12 months in the Wuhan cohort.[5] In our clusters, female sex and obesity were also associated with more severe ongoing health impairments including reduced exercise performance and HRQoL at one year, potentially highlighting a group which may need higher intensity interventions such as supervised rehabilitation. Health-related quality of life pre-COVID was substantially greater than at five-months post-discharge across all four clusters indicating that the persistent burden of impaired physical and mental health is not simply explained by pre-existing morbidity. The total number and range of ongoing symptoms at one year was striking, positively associated with the severity of Long-COVID, and emphasises the multi-system nature of Long-COVID. Other studies have shown the number of symptoms during the acute illness were associated with the likelihood of developing Long-COVID.[29] Whether the number of ongoing symptoms, a simple widely available measure, could underpin a future risk score deserves further attention.

Taken together we suggest our data will help inform decisions about patient stratification for follow-up after hospital discharge. We advocate a proactive approach due to the high proportion of non-recovery, highlight the utility of a screening questionnaire of whether patients feel fully recovered, the total number of symptoms may be a guide to the intensity/complexity of care required, and similar to our five-month data[3], we highlight the need for an holistic assessment including mental health, physical function and cognitive impairment. Any assessment of ongoing organ impairment will need to be further individualised.

Currently, there are no specific therapeutics for Long-COVID and our data highlights effective interventions are urgently required. Critically, our findings of persistent systemic inflammation, particularly in those in the 'very severe' and 'moderate/cognitive' clusters, suggests these groups might respond to anti-inflammatory strategies. The upregulation of IL6 suggests that anti-IL6 biologics that were successful for acute-hospitalised COVID-19[30] might also have a place in Long-COVID. Similarly, activation of the uPAR pathway suggests that IL1 activation might play a role, with soluble uPAR a biomarker in acute COVID-19 associated with good response to the recombinant IL-1 receptor antagonist anakinra.[31] Impaired exercise capacity was also associated with the more severe clusters and showed minimal improvement at one-year (below the minimum clinically important distance for other long-term conditions).[32-34] Current therapies available for some adults with Long-COVID include rehabilitation,[35] but the optimal exercise prescription is contentious due to concerns of post-exertional symptom exacerbation (PESE). Our data confirm high prevalence of musculoskeletal symptoms including muscle ache, fatigue, breathlessness, physically slowing down, and limb weakness.[5,16] This supports the need to investigate rehabilitation in

427 combination with other therapies to improve skeletal muscle function such as mitochondrial
428 energetics, mitophagy enhancers and drugs to combat cell senescence (associated with ageing).
429
430 The concordance of the severity of physical and mental health impairment in Long-COVID
431 highlights the need for close integration between physical and mental healthcare for Long-
432 COVID including assessment and interventions, but also for healthcare professional
433 knowledge transfer to improve patient care. It also suggests the need for complex interventions
434 that target both physical and mental health impairments in order to ameliorate symptoms.
435 However, specific therapeutic approaches to manage post-traumatic stress disorder may be
436 needed.[36] With obesity being associated with both non-recovery and severity of Long-COVID,
437 whether weight reduction using pharmacological and non-pharmacological approaches in
438 concert can ameliorate Long-COVID warrants further investigation. Beyond diet and lifestyle
439 interventions, glucagon like peptide-1 (GLP-1) analogues have recently been reported to
440 achieve clinically important weight reduction.[37]
441
442 Our cohort study is ongoing, and we report these one-year findings to add novel findings to the
443 limited literature, and help direct clinical care and further investigation. However, there are
444 limitations. There will be selection bias for participants returning for a one-year visit, although
445 we have not found overt differences between the demographics, or five-month recovery status
446 between attendees and non-attendees of the one-year visit. . Our cohort has a higher proportion
447 of patients requiring IMV than typically seen in UK hospitals[38] and therefore our results may
448 not be directly generalisable to the wider population. We also have a lower than expected
449 proportion of females which might mean that the wider population have worse outcomes than
450 we report as females appear to have worse recovery. To reduce uncertainty of the impact of
451 pre-existing illness, we asked our recruits whether they felt fully recovered i.e., back to their

normal. We also asked them retrospectively to estimate their pre-COVID-19 health status including the most prevalent symptoms, disability and health-related quality of life; we recognise there might be recall bias. Data linkage to electronic patient records is in process, but not currently available so in the current report pre-existing co-morbidities were self-reported and data regarding hospital admissions and mortality in the first year are unavailable. Our study suggests that persistent inflammation may be underlying ongoing impairment in some participants; the specific mechanisms underlying this signal require further investigation and replication. We describe several associations with more severe health impairments at one year. Our findings cannot confirm causality but suggest these associations should be further investigated as part of mechanistic studies and clinical trials. Our results require interpretation in the context of the evolving pandemic. Our one-year findings involve patients discharged from hospital in 2020 and therefore would not include newer variants such as Omicron and would not have been vaccinated prior to contracting COVID-19. Although our data has direct relevance to the large legacy of patients discharged under similar conditions, further research is needed to understand the impact of current acute care, newer variants and vaccination status prior to and after contracting COVID-19.

In summary, our study highlights an urgent need for healthcare services to support this large and rapidly increasing patient population where, in our cohort, there is a substantial burden of symptoms, reduced exercise capacity and large decrements in health-related quality of life after one year. Without effective treatments, Long-COVID has the potential to become a highly prevalent new long-term condition. Our study also provides a rationale for investigating treatments for Long-COVID with a precision medicine approach to target treatments to the relevant phenotype = to restore health-related quality of life.

Contributors

The manuscript was initially drafted by RAE, CEB, LVW and further developed by the writing committee. CEB, RAE, LVW, OE, HJCM, AShi, ASi, MJD, ABD, NIL, AShe, JDC, L-PH, AH, MM, KP, BR, made substantial contributions to the conception and design of the work. RAE, ASi, MS, RMS, VCH, RA, PB, CEBo, JSB, GC, NDB, NE, CE, JF, NH, JRH, MGJ, DP, PP, NMR, SLR-J, AMS, DGW, JDC, L-PH, AH, MM, WD-CM made substantial contributions to the acquisition of data. CEB, RAE, LVW, OCL, MR, OE, HJCM, MS, TC, MJD, ADS, JRG, WG, NJG, LGH, SH, LSH, JJ, RGJ, JML, WD-CM, GPM, SN, PJMO, JP, JQ, MJR, JTS, MGS, SJS, MTo, KEL, RST, AB, ABD, SK, NIL, AShe, MTh, BZ, JDC, L-PH, AH, MM, KP, BR, EMH, made contributions to the analysis, or interpretation of data for the work. All authors contributed to data interpretation and critical review and revision of the manuscript. Final approval of the version to be published and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. RAE, HJCM, OCL have accessed and verified the data. RAE, CEB and LVW were responsible for the decision to submit the manuscript.

Writing Group (on behalf of the PHOSP-COVID Collaborative Group†)

Rachael A Evans* [1], Olivia C Leavy [2], Matthew Richardson [1], Omer Elneima [1], Hamish J C McAuley [1], Aarti Shikotra [3], Amisha Singapuri [1], Marco Sereno [1], Ruth M Saunders [1], Victoria C Harris [1,4], Linzy Houchen-Wolloff [5,6,7], Raminder Aul [8], Paul Beirne [9], Charlotte E Bolton [10,11,12], Jeremy S Brown [13], Gourab Choudhury [14,15], Nawar Diar Bakerly [16,17], Nicholas Easom [18,19], Carlos Echevarria [20,21], Jonathan Fuld [22,23,24], Nick Hart [25,26], John R Hurst [27,28], Mark G Jones [29,30], Dhruv Parekh [31,32], Paul Pfeffer [33,34], Najib M Rahman [35,36,37,38], Sarah L Rowland-Jones [39,40], Ajay M Shah[41,42], Dan G Wootton [43,44,45], Trudie Chalder [46,47], Melanie J Davies [3,4,48], Anthony De Soyza [49,50], John R Geddes [51,52], William Greenhalf [44,53,54], Neil J Greening [1], Liam G Heaney [55,56], Simon Heller [57], Luke S Howard [58,59], Joseph Jacob [60,61], R Gisli Jenkins [62], Janet M Lord [63,64], William D-C Man [62,65], Gerry P McCann [3,66], Stefan Neubauer [36,37, 38,67] , Peter JM Openshaw [62], Joanna C Porter [68,69], Matthew J Rowland [70], Janet T Scott [71], Malcolm G Semple [72,73], Sally J Singh [1,2,3,4,5,6], David Thomas [59], Mark Toshner [23, 24,74], Keir E Lewis [75,76,77], Ryan S Thwaites [62], Andrew Briggs [78], Annemarie B Docherty [79], Steven Kerr [79,80], Nazir I Lone [81,82], Jennifer Quint [62], Aziz Sheikh [81], Mathew Thorpe [79], Bang Zheng [14], James D Chalmers [83], Ling-Pei Ho [35,37,84] , Alex Horsley [85,86], Michael Marks [87,88,89], Krisnah Poinasamy [90], Betty Raman [35,67], Ewen M Harrison [79], Louise V Wain* [1,2], Christopher E Brightling* [1]

*Joint first and last authors and contributed equally

†Details of the PHOSP-COVID Collaborative Group membership is provided as a supplementary file

Author Institutions

1. The Institute for Lung Health, NIHR Leicester Biomedical Research Centre, University of Leicester, Leicester, UK
2. Department of Health Sciences, University of Leicester, Leicester, UK
3. NIHR Leicester Biomedical Research Centre, University of Leicester, Leicester, UK

4. University Hospitals of Leicester NHS Trust, Leicester, UK
5. Centre for Exercise and Rehabilitation Science, NIHR Leicester Biomedical Research Centre-Respiratory, University of Leicester, Leicester, UK
6. Department of Respiratory Sciences, University of Leicester, Leicester, UK
7. Therapy Department, University Hospitals of Leicester NHS Trust, Leicester, UK
8. St Georges University Hospitals NHS Foundation Trust, London, UK
9. The Leeds Teaching Hospitals NHS Trust, Leeds, UK
10. University of Nottingham, Nottingham, UK
11. Nottingham University Hospitals NHS Trust, Nottingham, UK
12. NIHR Nottingham Biomedical Research Centre, Nottingham, UK
13. UCL Respiratory, Department of Medicine, University College London, Rayne Institute, London, UK
14. University of Edinburgh, Edinburgh, UK
15. NHS Lothian, Edinburgh, UK
16. Manchester Metropolitan University, Manchester, UK
17. Salford Royal NHS Foundation Trust, Manchester, UK
18. Infection Research Group, Hull University Teaching Hospitals, Hull, UK
19. University of Hull, Hull, UK
20. The Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle Upon Tyne, UK
21. Translational and Clinical Research Institute, Newcastle University, Newcastle Upon Tyne, UK
22. Department of Respiratory Medicine, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
23. University of Cambridge, Cambridge, UK
24. NIHR Cambridge Clinical Research Facility, Cambridge, UK
25. Lane Fox Respiratory Service, Guys and St Thomas NHS Foundation Trust, London, UK
26. Lane Fox Clinical Respiratory Physiology Research Centre, Centre for Human Applied Physiological Science, King's College London
27. University College London, London, UK
28. Royal Free London NHS Foundation Trust, London, UK
29. Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK
30. NIHR Southampton Biomedical Research Centre, University Hospitals Southampton, Southampton, UK
31. Birmingham Acute Care Research Group, University of Birmingham, Birmingham, UK
32. University Hospital Birmingham NHS Foundation Trust, Birmingham, UK
33. Barts Health NHS Trust, London, UK
34. Queen Mary University of London, London, UK
35. Oxford University Hospitals NHS Foundation Trust, Oxford, UK
36. University of Oxford, Oxford, UK
37. NIHR Oxford Biomedical Research Centre, Oxford, UK
38. CAMS Oxford Institute, Oxford, UK
39. University of Sheffield, Sheffield, UK
40. Sheffield Teaching NHS Foundation Trust, Sheffield, UK
41. King's College London, British Heart Foundation Centre, London, UK
42. King's College Hospital NHS Foundation Trust, London, UK
43. Institute of Infection, Veterinary and Ecological Sciences, University of

Liverpool, Liverpool, UK
44. Liverpool University Hospitals NHS Foundation Trust, Liverpool, UK
45. NIHR Health Protection Research Unit in Emerging and Zoonotic Infections, University of Liverpool, Liverpool, UK
46. Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, UK
47. South London and Maudsley NHS Trust, London, UK
48. Diabetes Research Centre, University of Leicester, Leicester, UK
49. Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK
50. Newcastle upon Tyne Teaching Hospitals Trust, Newcastle upon Tyne, UK
51. NIHR Oxford Health Biomedical Research Centre, University of Oxford, Oxford, UK
52. Oxford Health NHS Foundation Trust, Oxford, UK
53. University of Liverpool, Liverpool, UK
54. The CRUK Liverpool Experimental Cancer Medicine Centre, Liverpool, UK
55. Wellcome-Wolfson Institute for Experimental Medicine, Queens University Belfast, Belfast, UK
56. Belfast Health & Social Care Trust, Belfast, UK
57. Department of Oncology and Metabolism, University of Sheffield, Sheffield, UK
58. Imperial College Healthcare NHS Trust, London, UK
59. Imperial College London, London, UK
60. Centre for Medical Image Computing, University College London, London, UK
61. Lungs for Living Research Centre, University College London, London, UK
62. National Heart and Lung Institute, Imperial College London, London, UK
63. MRC-Versus Arthritis Centre for Musculoskeletal Ageing Research, Institute of Inflammation and Ageing, University of Birmingham, Birmingham, UK
64. NIHR Birmingham Biomedical Research Centre, University Hospitals Birmingham and the University of Birmingham, Birmingham, UK
65. Royal Brompton and Harefield Clinical Group, Guy's and St Thomas' NHS Foundation Trust, London, UK
66. Department of Cardiovascular Sciences, University of Leicester, Leicester, UK
67. Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK
68. UCL Respiratory, Department of Medicine, University College London, Rayne Institute, London, UK
69. ILD Service, University College London Hospital, London, UK
70. Kadoorie Centre for Critical Care Research, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK
71. MRC-University of Glasgow Centre for Virus Research, Glasgow, UK
72. NIHR Health Protection Research Unit in Emerging and Zoonotic Infections, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, UK
73. Respiratory Medicine, Alder Hey Children's Hospital, Liverpool, UK
74. Cambridge NIHR BRC, Cambridge, UK
75. Hywel Dda University Health Board, Wales, UK
76. University of Swansea, Wales, UK
77. Respiratory Innovation Wales, Wales, UK
78. London School of Hygiene & Tropical Medicine, London, UK
79. Centre for Medical Informatics, The Usher Institute, University of Edinburgh,

Edinburgh, UK
80. Roslin Institute, University of Edinburgh, Edinburgh, UK
81. Usher Institute, University of Edinburgh, Edinburgh, UK
82. Royal Infirmary of Edinburgh, NHS Lothian, Edinburgh, UK
83. University of Dundee, Ninewells Hospital and Medical School, Dundee, UK
84. MRC Human Immunology Unit, University of Oxford, Oxford, UK
85. Division of Infection, Immunity & Respiratory Medicine, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
86. Manchester University NHS Foundation Trust, Manchester, UK
87. Department of Clinical Research, London School of Hygiene & Tropical Medicine, London, UK
88. Hospital for Tropical Diseases, University College London Hospital, London, UK
89. Division of Infection and Immunity, University College London, London, UK
90. Asthma + Lung UK, London, UK, London, UK

Acknowledgments

This study would not be possible without all the participants who have given their time and support. We thank all the participants and their families. We thank the many research administrators, health-care and social-care professionals who contributed to setting up and delivering the study at all of the 69 NHS Trusts and 25 Research Institutions across the UK as well as all the supporting staff at the at the NIHR Clinical Research Network, Health Research Authority, Research Ethics Committee, Department of Health and Social Care Public Health Scotland, Public Health England and support from the ISARIC Coronavirus Clinical Characterisation Consortium (ISARIC4C). At the NIHR Office for Clinical Research Infrastructure (NOCRI) we thank Kate Holmes (for her support in coordinating the charities group). We are very grateful to all the charities that have provided insight to the study- Action Pulmonary Fibrosis, Alzheimer's Research UK, Asthma UK /British Lung Foundation UK/BLF, British Heart Foundation, Diabetes UK, Cystic Fibrosis Trust, Kidney Research UK, MQ Mental Health, Muscular Dystrophy UK, Stroke Association Blood Cancer UK, McPin Foundations, Versus Arthritis. We thank the NIHR Leicester Biomedical Research Centre patient and public involvement group and the Long Covid Support Group. This research used the SPECTRE High Performance Computing Facility at the University of Leicester.

Data sharing

The protocol, consent form, definition and derivation of clinical characteristics and outcomes, training materials, regulatory documents, requests for data access and other relevant study materials are available online at www.phosp.org.

Declaration of interests

AB declares that their institute was awarded a grant from the NIHR to complete this work and receives consulting fees from Roche, Merck and GSK.

ADS declares grants from AstraZeneca, GlaxoSmithKline, Novartis, Sanofi and personal fees from 30T, LifeARC, Gilead, AstraZeneca. Payments for lectures from AZ, GSK, Gilead and Pfizer. Participation on a board for Bayer and a Trustee for Action for Pulmonary Fibrosis.

AH declares that their institute was awarded a grant from the UKRI/NIHR to complete this work, a grant from NIHR Manchester CRF to support study delivery and NIHR Manchester BRC for personal funding and institutional payments to support grant funded research from NIHR, MRC, Cystic Fibrosis Trust, Cystic Fibrosis Foundation, North West Lung Centre Charity and Moulton Trust. Consulting fees from Mylan Pharmaceuticals for advisory board participation. Payment from Vertex Pharmaceuticals for educational presentation, participation on a clinical trials advisory board and writing review article. Non paid roles include; chair of Cystic Fibrosis Clinical Trials Accelerator Program, deputy chair NIHR Respiratory Translational Research Collaboration and director of University spin out company (Mi-trial Ltd).

AMS declares that their institute was awarded a grant from the UKRI/NIHR to complete this work and research grants from British Heart Foundation, MRC and NIHR-BRC.

AShe declares that their institute was awarded a grant from the UKRI. They declare unremunerated participation on AstraZeneca Thrombotic Thrombocytopenic Taskforce and Scottish and UK Governments – COVID-19 Advisory Groups

BR declares that payments are received from British Heart Foundation Oxford Centre of Research Excellence, NIHR Oxford BRC and UKRI for grants and contracts. Consulting fees have been received from Axcella Therapeutics.

CE declares funding from GlaxoSmithKline for investigator led research project.

CEB – declares that their institute was awarded a grant from the UKRI/NIHR to complete this work. Grants from GSK, AZ, Sanofi, BI, Chiesi, Novartis, Roche, Genentech, Mologic, 4DPharma. Consultancy paid to institution from GSK, AZ, Sanofi, BI, Chiesi, Novartis, Roche, Genentech, Mologic, 4DPharma, TEVA.

CEBo declares that their institute was awarded a grant from the UKRI/NIHR and institutional support from NIHR Nottingham BRC to complete this work. Grants to support post covid project from NIHR Nottingham BRC and NUH R&I and Nottingham Hospitals Charity.

DGW declares they are supported by an Advanced Fellowship from NIHR.

DP declares that their institute was awarded a grant from NIHR and MRC and hold leadership roles within the British Thoracic Society

GC declares that their institute was awarded a grant from GSK, Astra Zeneca, British Lung Foundation, Mereo and Arrowhead Pharmaceutials. Personal payments was received from GSK and Astra Zeneca for educational meetings and presentations. GSK paid for conference registration fees. Unpaid participation as chair of the Lothian respiratory MCN and Act on COPD Group in Scotland for Astra Zeneca.

GPM declares that their institute was awarded a grant from the UKRI/NIHR to complete this work, grants from BHF and MRC, support for attending meetings from the British and Irish Society for Minimally Invasive Cardiac Surgery, leadership in the British Society for Cardiovascular MRI and receipt of research software from Circle CVi.

JRH declares consultancy fees from AstraZeneca, speaker fees from Boehringer Ingelheim, Takeda and travel grants from AstraZeneca. Participation on advisory board for AstraZeneca

and unpaid leadership role with the British Thoracic Society. Received donation of oximeters from Nonin.

JCP declares their institution received grants from UKRI, LifeArc and MRC. Payment fees from The Limbic and advisory board membership at Carrick Therapeutics and AstraZeneca.

JDC declares grants received from AstraZeneca, Boehringer Ingelheim, Insmed, Novartis, Gilead Sciences and Genentech, consulting fees from AstraZeneca, Boehringer Ingelheim, Insmed, Novartis, Gilead Science, Chiesi, Zambon and Genentech

JJ declares consulting fees from Boehringer Ingelheim, Roche, GlaxoSmithKline, NHSX and speaker fees from Boehringer Ingelheim, Roche, GlaxoSmithKline, Takeda and support to attend meeting from Boehringer Ingelheim. Participation on advisory boards at Boehringer Ingelheim and Roche. UK patent application number 2113765.8.

LH-W declares a grant from NIHR unrelated to the submitted work. Independent chair of the NIHR HTA Committee for Colour COPD trial and member of American Thoracic Society Pulmonary Rehabilitation Assembly Web and Planning Committees.

L-PH declares their institution received grants from UKRI, Regenerative Medicine Platform, Celgene, British Lung Foundation and Oxford Boehnringer Ingleheim. Advisory Board for CATALYST trial and Chair of Respiratory Translational Research Collaboration.

LVW declares research funding unrelated to the submitted work from GSK and Orion, consulting fees unrelated to the submitted work from Galapgos and Wellcome Conference speaker honorarium, travel support from Genentech, participation in advisory board for Galapgos, an associate editor role for European Respiratory Journal.

MGJ declares that their institute was awarded a grant from Boehringer Ingelheim

MGS declares that their institute was awarded a grant from the NIHR, MRC UK and Health Protection Research Unit in Emerging & Zoonotic Infections, University of Liverpool to complete this work, Independent external and non-remunerated member of Pfizer's External Data Monitoring Committee for their mRNA vaccine program, Chair of Infectious Disease Scientific Advisory Board for Integrum Scientific LLC, minority share owner Integrum Scientific LLC, and a non-remunerated independent member of HMG Scientific Group for Emergencies (SAGE) and HMG UK New Emerging Respiratory Virus Threats Advisory Group (NERVTAG).

MJD declares payments to their institution from AstraZeneca, Novo Nordisk, Boehringer Ingelheim and Janssen outside the submitted work. Consulting fees from Novo Nordisk, Eli Lilly and Boehringer Ingelheim. Personal fees for lectures and presentations from Novo Nordisk, Sanofi-Aventis, Eli Lilly, Boehringer Ingelheim, AstraZeneca and Napp Pharmaceuticals, a member of RESiliENT Trial Steering Committee and Chair of EASD Writing Group.

MJR declares a grant from NIHR for HTA SOS Trial and NIHE EME Programme study (OSMOTIC) and current employment by Roche Pharmaceuticals on a one year academic/industry senior clinical fellowship.

NE received donations of COVID-19 lateral flow tests for a pilot project from Mologics Ltd

NIL declares Director of Research, Intensive Care Society UK

PJMO declares co-funding from the MRC and GSK (INFLAMMAGE), part of the EMINENT consortium to promote inflammation research. Consulting fees from Jansssen, Sequiris, Valneva SE, payments for speaking from Janssen and Seqirus, member and vice-chair of NERVTAG.

PP declares grants from the NIHR to the institute to support remote rehabilitation post Covid-19.

RGJ declares a commercial contract with PatientMPower to provide App and spirometers with no payments by PatientMPower for the study. Payments to their institution from Astra Zeneca, Biogen, Galecto, GlaxoSmithKline, RedX and Pliant. Consulting fees from Bristol Myers Squibb, Daewoong, Veracyte, Resolution Therapeutics and Pliant. Payments for lectures form Chiesi, Roche, PatientMPower and AstraZeneca. Participation on advisory boards at Boehringer Ingelheim, Galapagos and Vicore. Leadership role at NuMedii and a trustee for Action for Pulmonary Fibrosis.

RAE declares that their institute was awarded a grant from the UKRI/NIHR to complete this work. Speaker fees from Boehringer Ingelheim and unpaid roles of European Respiratory Society Assembly 01.02 Pulmonary Rehabilitation secretary and American Thoracic Society Pulmonary Rehabilitation Assembly programme committee.

SH declares grants from the European Commission and NIHR. Consulting fees from Eli Lilly, Zealand Pharma, NovoNordisk and Mylan. Honorary payments from NovoNordisk and payment for expert testimony from the Crown Prosecution Service.

SN declares research grant from Axcella.

SLR-J declares that their institute was awarded a grant from the UKRI/NIHR and salary support from CRN to complete this work. Grants from UKRI, NIHR, GCRF and EDCTP for unrelated studies. Participation on Data Safety Monitoring Board for 2 trials (Bexero for gonococcal infection in Kenya, inactivated Covid vaccine trial in Zimbabwe) and previous president of Royal Society of Tropical Medicine and Hygiene.

TC declares that their institute was awarded a grant from NIHR BRC at South London and Maudsley NHS Foundation Trust to complete this work, grant from NIHR for CLOCK study. Speaker fees from Hello Self, BABCP and Department of Health. Unpaid participation on NICE guideline committee on Post/ Long COVID. Leads Persistent Physical Symptom service as part of their paid employment. Author of a published self-help books on fatigue for which received payments.

WD-CM declares grants from NIHR, British Lung Foundation and NHSX, payments for lectures from Munipharma, Novartis and European Conference and Incentive Services DMC. Participated on a monitoring board for Jazz Pharmaceuticals. Received funds for blood analysis from GSK.

AShi, ASi, JML, MM and **NDB** declare that their institute was awarded a grant from the UKRI/NIHR to complete this work. All other authors declare no competing interests.

Funding

PHOSP-COVID is supported by a grant from the MRC-UK Research and Innovation and the Department of Health and Social Care through the National Institute for Health Research (NIHR) rapid response panel to tackle COVID-19 (grant references: MR/V027859/1 and COV0319). Core funding was provided by NIHR Leicester Biomedical Research Centre to support the PHOSP-COVID coordination team and NIHR Biomedical Research Centres (BRCs), Clinical Research Facilities (CRF) and NIHR Health Protection Research Unit (HPRU) and Translational Research Collaborations (TRCs) network across the country. The institutional funding that supports the outbreak labs that process the PHOSP samples NIHR Health Protection Research Unit (HPRU) in Emerging and Zoonotic Infections at University of Liverpool in partnership with Public Health England (PHE) and Liverpool Experimental Cancer Medicine Centre (grant reference: C18616/A25153)

ABD is funded by a Wellcome Trust grant (216606/Z/19/Z)
RAE* holds a National Institute for Health Research (NIHR) clinician scientist fellowship (CS-2016-16-020).
NJG holds a NIHR post-doctoral fellowship (PDF- 2017- 10- 052).
JJ was supported by a Wellcome Trust Clinical Research Career Development Fellowship 209553/Z/17/Z] and by the NIHR University College London Hospital Biomedical Research Centre, UK.
LVW* was supported by a GSK / British Lung Foundation Chair in Respiratory Research (C17-1)
DGW is supported by an NIHR Advanced Fellowship NIHR300669

Funder role

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

The views expressed in this publication are those of the author(s) and not necessarily those of the MRC, NIHR or the Department of Health and Social Care. No form of payment was given to anyone to produce the manuscript. All members of the writing group have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest.

Ethics Approval Ethics Ref: 20/YH/0225 Trial ID: ISRCTN10980107

References

1. Coronavirus Resource Centre – John Hopkins University. https://coronavirus.jhu.edu/map.html. Accessed 6 December 2021

2. UK summary coronavirus (COVID-19) in the UK. https://coronavirus.data.gov.uk. Accessed 6 December 2021

3. Evans RA, McAuley H, Harrison EM et al. Physical, cognitive, and mental health impacts of COVID-19 after hospitalisation (PHOSP-COVID): a UK multicentre, prospective cohort study. *Lancet Respir Med*. 2021 Nov; **9**(11):1275-1287

4. 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.

5. Huang L, Yao Q, Gu X et al. 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. *Lancet* 2021; **398**(10302):747-758.

6. Filbin MR, Mehta A, Schneider AM, et al. Longitudinal proteomic analysis of severe COVID-19 reveals survival-associated signatures, tissue-specific cell death, and cell-cell interactions. *Cell Rep Med*. 2021; **2**(5):100287.

7. Thwaites RS, Sanchez Sevilla Uruchurtu A, Siggins MK, et al. Inflammatory profiles across the spectrum of disease reveal a distinct role for GM-CSF in severe COVID-19. Sci Immunol. 2021; 6(57):eabg9873.

8. COVID-19 rapid guideline: managing the long-term effects of COVID-19: https://www.nice.org.uk/guidance/ng188, updated November, 2021. Accessed 12.12.2021

9. A clinical case definition of post COVID-19 condition by a Delphi consensus https://www.who.int/publications/i/item/WHO-2019-nCoV-Post_COVID-19_condition-Clinical_case_definition-2021.1 6 October 2021. Accessed 12.12.2021

10. Kaufman L, Rousseeuw PJ. Clustering large applications (Program CLARA). *Finding groups in data: an introduction to cluster analysis* 2008: 126-63.

11. Fu W, Liu Y, Liu L, et al. An open-label, randomized trial of the combination of IFN-κ plus TFF2 with standard care in the treatment of patients with moderate COVID-19. *EClinicalMedicine*. 2020;27:100547.

12. Tserel L, Jõgi P, Naaber P, et al. Long-Term Elevated Inflammatory Protein Levels in Asymptomatic SARS-CoV-2 Infected Individuals. *Front Immunol*. 2021;12:709759.

13. Fraser DD, Patterson EK, Slessarev M, et al. Endothelial Injury and Glycocalyx Degradation in Critically Ill Coronavirus Disease 2019 Patients: Implications for Microvascular Platelet Aggregation. *Crit Care Explor*. 2020;2(9):e0194.

14. Laudanski K, Jihane H, Antalosky B, et al. Unbiased Analysis of Temporal Changes in Immune Serum Markers in Acute COVID-19 Infection With Emphasis on Organ Failure, Anti-Viral Treatment, and Demographic Characteristics. *Front Immunol*. 2021;12:650465.

15. Hettwer S, Dahinden P, Kucsera S, et al. Elevated levels of a C-terminal agrin fragment identifies a new subset of sarcopenia patients. *Exp Gerontol*. 2013;48(1):69-75.

16. Dhaeze, T., Tremblay, L., Lachance, C. *et al*. CD70 defines a subset of proinflammatory and CNS-pathogenic T_H1/T_H17 lymphocytes and is overexpressed in multiple sclerosis. *Cell Mol Immunol* **16,** 652–665 (2019)

17. Griffith DM, Vale ME, Campbell C, Lewis S, Walsh TS. Persistent inflammation and recovery after intensive care: A systematic review. *J Crit Care*. 2016;33:192-199.

18. Fernández-de-Las-Peñas C, Guijarro C, Plaza-Canteli S, Hernández-Barrera V, Torres-Macho J. Prevalence of Post-COVID-19 Cough One Year After SARS-CoV-2 Infection: A Multicenter Study. *Lung*. 2021;199(3):249-253.

19. 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.

20. Latronico N, Peli E, Calza S, et al. Physical, cognitive and mental health outcomes in 1-year survivors of COVID-19-associated ARDS [published online ahead of print, 2021 Sep 29]. *Thorax*. 2021;thoraxjnl-2021-218064.

21. Dowdy DW, Eid MP, Dennison CR, et al. Quality of life after acute respiratory distress syndrome: a meta-analysis. *Intensive Care Med*. 2006;32(8):1115-1124.

22. Parry SM, Nalamalapu SR, Nunna K, et al. Six-Minute Walk Distance After Critical Illness: A Systematic Review and Meta-Analysis. *J Intensive Care Med*. 2021;36(3):343-

23. Whitaker M, Elliott J, Chadeau-Hyam M et al. Persistent symptoms following SARS-CoV-2 infection in a random community sample of 508,707 people medRxiv 2021.06.28.21259452; doi: https://doi.org/10.1101/2021.06.28.21259452

24. Blomberg B, Mohn KG, Brokstad KA, et al. Long COVID in a prospective cohort of home-isolated patients. *Nat Med*. 2021;27(9):1607-1613

25. Wynberg E, van Willigen HDG, Dijkstra M, et al. Evolution of COVID-19 symptoms during the first 12 months after illness onset [published online ahead of print, 2021 Sep 2]. Clin Infect Dis. 2021;ciab759.

26. Thompson EJ, Williams DM, Walker AJ et al. Risk factors for long COVID: analyses of 10 longitudinal studies and electronic health records in the UK. medRxiv 2021.06.24.21259277; doi: https://doi.org/10.1101/2021.06.24.21259277

27. UK Office of National Statistics COVID-19 https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptomsfollowingcoronaviruscovid19infectionintheuk/2december2021. Accessed 6 December, 2021.

28. Boscolo-Rizzo, P, Guida, F, Polesel, J, et al. Sequelae in adults at 12 months after mild-to-moderate coronavirus disease 2019 (COVID-19). *Int Forum Allergy Rhinol*. 2021; 11: 1685–1688.20 (PMC296319).

29. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. *Nat Med*. 2021; **27**(4):626-631.

30. Ghosn L, Chaimani A, Evrenoglou T, et al. Interleukin-6 blocking agents for treating COVID-19: a living systematic review. *Cochrane Database Syst Rev*. 2021; **3**(3):CD013881.

31. Kyriazopoulou E, Poulakou G, Milionis H, et al. Early treatment of COVID-19 with anakinra guided by soluble urokinase plasminogen receptor plasma levels: a double-blind, randomized controlled phase 3 trial. *Nat Med*. 2021; **27**(10):1752-1760.

32. Evans RA, Singh SJ. Minimum important difference of the incremental shuttle walk test distance in patients with COPD. *Thorax*. 2019; **74**(10):994-995.

33. Nolan CM, Delogu V, Maddocks M, et al. Validity, responsiveness and minimum clinically important difference of the incremental shuttle walk in idiopathic pulmonary fibrosis: a prospective study. *Thorax* 2018*;* **73**(7):680-682.

34. Parreira VF, Janaudis-Ferreira T, Evans RA, Mathur S, Goldstein RS, Brooks D. Measurement properties of the incremental shuttle walk test, a systematic review. *Chest*. 2014; **145**(6):1357-1369. doi:10.1378/chest.13-2071

35. Spruit MA, Holland AE, Singh SJ, Tonia T, Wilson KC, Troosters T. COVID-19: Interim Guidance on Rehabilitation in the Hospital and Post-Hospital Phase from a European Respiratory Society and American Thoracic Society-coordinated International Task Force. *Eur Respir J*. 2020; **56**(6):2002197. doi:10.1183/13993003.02197-2020

36. Martin A, Naunton M, Kosari S, Peterson G, Thomas J, Christenson JK. Treatment Guidelines for PTSD: A Systematic Review. *J Clin Med*. 2021; **10**(18):4175.

37. Wilding JPH, Batterham RL, Calanna S, et al. Once-Weekly Semaglutide in Adults with Overweight or Obesity. *N Engl J Med*. 2021; **384**(11):989.

38. Docherty AB, Mulholland RH, Lone NI, et al. Changes in in-hospital mortality in the first wave of COVID-19: a multicentre prospective observational cohort study using the WHO Clinical Characterisation Protocol UK. *Lancet Respir Med*. 2021;9(7):773-785

Tables and Figures

Main manuscript

Tables

Table 1. Individual and hospital admission characteristics for the participants who had a five-month visit, a one-year visit, and both visits.

Table 2. A comparison of the change between five-months and one-year post-discharge in patient reported outcome measures including mental health, physical function, cognitive impairment, and organ function

Table 3. A comparison of the change in patient reported outcome measures including mental health, physical function, cognitive impairment, and organ function between five-month and one-year sub-grouped by the four 'severity' cluster phenotypes

Figures

Figure 1. Consort diagram of participants

Figure 2. Four panel plot demonstrating patients perceived recovery at one year a) compared with five months, b) risk factors for being less likely to recover, c) compared by the four clusters d) compared to health-related quality of life (assessed by the EQ5D-5L Utility Index)

Figure 3. Volcano plots representing multinomial regression association results for comparison of 368 proteins between a) clusters 1 (very severe) and 4 (mild), b) clusters 2 (severe) and 4 (mild) and c) clusters 3 (moderate/cognitive) and 4 (mild). The red horizontal line represents an unadjusted $P<0{\cdot}05$ threshold. Proteins that were significantly over-expressed (compared to the reference cluster 4) after FDR adjustment are indicated in red and those significantly under-expressed are indicated in black, FDR cut off used was $0{\cdot}1$.

Figure 4. **Figure 4. Characteristics associated with the four recovery clusters**

a) Patient characteristics, CRP level, exercise performance and symptom count across the four clusters

b) Health-related quality of life across the four clusters assessed pre-hospitalisation (patient estimate), and at five-months and one-year post-discharge

Supplement Tables and Figures

Table S1. Tier 2 outcome measures.

Table S2. Methods and thresholds for processing of variables and outcome measures used in the current analysis.

Table S3. Comparison of individual and admission characteristics, and patient-perceived recovery between the one-year visit attendees and non-attendees for participants discharged between 28th February 2020 – 28th November June 2020.

Table S4. Comparison between imputed and non-imputed logistic regression of predictors of failure to recover at 1-year (multi-variable and multi-level).

Table S5. Patient reported outcomes, physiological and biochemical tests stratified by patient perceived recovery outcomes at five-months and one-year visits a) paired data and b) unpaired data.

Table S6. Cluster medoids and characteristics

Table S7. Comparison of participant demographics, clinical characteristics, and admission characteristics stratified by the four clinical recovery clusters.

Table S8: Multinomial logistic regression of clinical characteristics against cluster membership (mild cluster as reference) for participants with plasma proteomic data.

Table S9. Comparison of plasma proteomics between a) cluster 1 (very severe) and cluster 4 (mild), b) cluster 2 (severe) and cluster 4 (mild) and c) cluster 3 (moderate/cognitive) and cluster 4 (mild). Multinomial regression results with unadjusted $P<0{\cdot}05$ are presented.

Table S10. Ongoing symptoms recorded at five-months and one-year post-discharge from hospital

Table S11. A comparison between the four clinical recovery phenotypes of health-related quality of life, disability, fatigue and breathlessness across pre-hospitalisation, and five-months and one-year post-discharge

Figure S1. A comparison of participant-perceived recovery between five months and one-year

Figure S2. Clusters of mental, cognitive, and physical health impairments at five-months

Figure S3. Estimation plots for the features significantly upregulated when comparing cluster 1 (very severe) to cluster 4 (mild) (panels A-M) and when comparing cluster 3 (moderate/cognitive) to cluster 4 (mild) (panels N and O)

994 Figure S4. Disability and symptoms across the four ‘recovery’ clusters assessed for pre-
995 hospitalisation (patient estimated), and at five-months and one-year post-discharge

996 Figure S5. Health-related quality of life, disability, and symptoms across the four ‘recovery’
997 clusters assessed for pre-hospitalisation (patient estimated), and at five-months and one-year
998 post-discharge (individual complete data for all three time-points

999

Figure 1. Consort diagram of participants

Figure 1 Legend. GAD-7 = Generalized Anxiety Disorder 7-item scale. PHQ-9 = Patient Health Questionnaire-9. PCL-5 Post Traumatic Stress Disorder Checklist. FACIT Fatigue Scale = Functional Assessment of Chronic Illness Therapy Fatigue Scale. SPPB=short physical performance battery. MoCA = Montreal Cognitive Assessment. WG-SS-SCo = Washington Group Short Set of Functioning Severity Continuum. VAS = Visual Analogue Scale

Figure 2. Four panel plot demonstrating patients perceived recovery at one year a) compared with five months, b) risk factors for being less likely to recover, c) compared by the four clusters d) compared to health-related quality of life (assessed by the EQ5D-5L Utility Index)

Figure 2 Legend.

BMI: Body Mass Index, WHO: World Health Organisation, WHO clinical progression scale classes are as follows: 3–4 = no continuous supplemental oxygen needed; 5 = continuous supplemental oxygen only; 6 = continuous positive airway pressure or bi-level positive pressure ventilation or high-flow nasal oxygen; and 7–9 = invasive mechanical ventilation or other organ support. The forest plot of the patient and admission characteristics associated with patient perceived recovery at one-year used multi-variable logistic regression and multiple imputation. EQ5D-5L pre-COVID was retrospectively completed by participants.

Figure 3. Volcano plots representing multinomial regression association results for comparison of 296 proteins between the four clinical phenotypes

Figure 3 Legend. Volcano plots representing multinomial regression association results, corrected for age, BMI and number of comorbidities, comparing 296 proteins between a) clusters 1 (very severe) and 4 (mild), b) cluster 2 (severe) and cluster 4 (mild) and c) cluster 3 (moderate/cognitive) and cluster 4 (mild). The red horizontal line represents an unadjusted P<0·05 threshold. Proteins that were significantly differentially-expressed (compared to the reference cluster 4) after False Detection Rate (FDR) adjustment are indicated in red, FDR cut off used was 0·1.

Figure 4. Characteristics associated with the four recovery clusters

a) Patient characteristics, CRP level, exercise performance and symptom count across the four clusters

b) Health-related quality of life across the four clusters assessed pre-hospitalisation (patient estimate), and at five-months and one-year post-discharge

Figure 4 Legend. Cluster 1 Red - 'very severe' physical and mental health impairment
Cluster 2 Blue - 'severe' physical and mental health impairment
Cluster 3 Green - 'moderate/cognitive' physical impairment and cognitive impairment
Cluster 4 Purple - 'mild' physical and mental health impairment
BMI: Body Mass Index, CRP: C-reactive protein assessed at one year, mean : Incremental Shuttle Walk Test distance % predicted assessed at one year, IMV: invasive mechanical ventilation, *median number of symptoms at one year.
EQ5D-5L utility index stratified by cluster, pre-hospital health status assessed retrospectively

Main Manuscript Tables

Table 1. Individual and hospital admission characteristics for the participants who had a five-month visit, a one-year visit, and both visits.

Table 2. A comparison of the change between five-months and one-year post-discharge in patient reported outcome measures including mental health, physical function, cognitive impairment, and organ function

Table 3. A comparison of the change in patient reported outcome measures including mental health, physical function, cognitive impairment, and organ function between five-month and one-year sub-grouped by the four 'severity' cluster phenotypes

Table 1. Individual and hospital admission characteristics for the participants who had a five-month visit, a one-year visit, and both visits.

	Complete five-month visit	Complete one-year visit	Completed both visits
Total N (%)	2320 (100·0)	924 (100·0)	807 (100·0)
Age at admission, years*	58·0 (12·6)	58·9 (12·5)	58·7 (12·5)
Female Sex at birth¶	855 (39·0)	319 (35·8)	279 (35·6)
Missing data	127	33	23
Ethnicity¶			
White	1685 (75·4)	681 (74·7)	596 (74·5)
South Asian	262 (11·7)	102 (11·2)	94 (11·8)
Black	154 (6·9)	68 (7·5)	57 (7·1)
Mixed	46 (2·1)	19 (2·1)	18 (2·2)
Other	87 (3·9)	42 (4·6)	35 (4·4)
Missing data	86	12	7
Index of Multiple Deprivation (IMD)¶			
1 (most deprived)	517 (22·6)	187 (20·4)	163 (20·4)
2	533 (23·3)	186 (20·3)	163 (20·4)
3	404 (17·7)	175 (19·1)	155 (19·4)
`4	396 (17·3)	160 (17·5)	137 (17·2)
5 (least deprived)	438 (19·1)	207 (22·6)	180 (22·6)
Missing data	32	9	9
BMI †	31·2 (27·7 - 36·1)	31·5 (27·7 - 35·8)	31·5 (27·7 - 35·7)
< 30 kg/m^2¶	840 (41·1)	349 (40·3)	316 (41·2)
≥ 30 kg/m^2¶	1204 (58·9)	517 (59·7)	451 (58·8)
Missing data	276	58	40
Smoking status¶			
Never/Ex-smoker/Current	1085 (54·7) / 864 (43·5)/ 36 (1·8)	429 (53·2)/ 369 (45·7)/ 9 (1·1)	350 (52·8)/ 301 (45·4)/ 12 (1·8)
Missing data	335	117	144
WHO clinical progression scale¶			
WHO – class 3-4	385 (16·9)	171 (18·6)	145 (18·0)
WHO – class 5	959 (42·2)	342 (37·1)	299 (37·1)
WHO – class 6	517 (22·7)	167 (18·1)	139 (17·2)
WHO – class 7-9	412 (18·1)	241 (26·2)	224 (27·8)
Missing data	47	< 5	0
Comorbidities¶			
Median number of comorbidities †	2·0 (0·0 to 3·0)	2·0 (0·0 to 3·0)	2·0 (0·0 to 3·0)
0	642 (27·7)	251 (27·2)	213 (26·4)
1	468 (20·2)	172 (18·6)	154 (19·1)
≥2	1210 (52·2)	501 (54·2)	440 (54·5)
Admission duration, days	13·9 (18·2)	17·0 (24·7)	17·8 (22·1)
Positive SARS-CoV-2 PCR¶	1916 (92·4)	796 (90·8)	700 (90·6)
Missing data	246	47	34
Systemic steroids¶	1173 (54·2)	251 (29·8)	226 (30·2)
Missing data	157	81	59

¶ = Number (%) with positive response, *Mean [SD], †Median [IQR]. Percentages are calculated by category after exclusion of missing data for that variable. N= Number. WHO classes are as follows: 3–4 = no continuous supplemental oxygen needed; 5 = continuous supplemental oxygen only; 6 = continuous or bi-level positive airway pressure ventilation or high-flow nasal oxygen; and 7–9 = invasive mechanical ventilation or other organ support. BMI = body-mass index. PCR= Polymerase Chain Reaction. §Therapeutic dose anticoagulation; does not include intermediate doses which were not recorded.

Table 2. Patient reported outcome measures, physical function and organ function at five-months stratified by patient-perceived recovery and compared at one-year after hospital discharge

a)

	Five-month visit – stratified by patient perceived recovery at five months, n=1965				Paired data at five-month and one-year visits, n=807			
	Recovered	**Not sure**	**Not recovered**	**P**	**N pairs with available data**	**Five-months**	**One-year**	**p**
Total n (%)	501 (25·5)	385 (19·6)	1079 (54·9)	N/A	**% recovered** n=590	151 (25·6)	168 (28·5)	0·12
Time to review from discharge, days †	166 (127- 191)	165 (122 - 191)	157 (119 - 189)	0·040	807	178 (156 - 197)	384 (359-409)	<0·0001
BMI †	29·4 (26·6-33·5)	31·5 (28·0-36·3)	31·6 (28·0-36·4)	<0·0001	602	30·7 (27·3-35·0)	31·1 (27·5-35·5)	<0·0001
< 30 kg/m^2	230 (54·5)	131 (40·2)	360 (38·8)	<0·0001	602	275 (45·7)	255 (42·4)	0·021
≥ 30 kg/m^2	192 (45·5)	195 (59·8)	568 (61·2)			327 (54·3)	347 (57·6)	
Median number of symptoms †	3 (1 - 7)	8 (4 - 15)	14 (8 - 20)	<0·0001	619	9 (4 - 16)	9 (4 - 17)	0·010
Fatigue VAS †	0·0 (0·0 - 2·0)	2·0 (0·0 - 5·0)	5·0 (2·0 - 8·0)	<0·0001	521	3·0 (0·0 - 6·0)	3·0 (0·0 - 6·0)	0·090
Breathlessness VAS †	0·0 (0·0 - 1·0)	1·5 (0·0 - 4·0)	4·0 (1·0 - 6·0)	<0·0001	524	2·0 (0·0 - 5·0)	2·0 (0·0 - 5·0)	0·052
Anxiety (GAD7 >8)¶	53 (11·3)	81 (22·6)	339 (33·4)	<0·0001	684	164 (24·0)	147 (21·5)	0·13
Depression (PHQ-9 ≥10)¶	47 (9·9)	96 (26·7)	426 (42·0)	<0·0001	680	181 (26·6)	169 (24·9)	0·25
PTSD (PCL-5 ≥38)¶	18 (3·8)	34 (9·5)	202 (20·0)	<0·0001	680	83 (12·2)	68 (10·0)	0·055
Dyspnoea-12 score*	2·1 (4·8)	5·1 (7·1)	8·9 (8·8)	<0·0001	702	6·0 (8·1)	5·5 (7·7)	0·040
FACIT fatigue subscale score*	43·6 (8·8)	36·5 (11·2)	29·1 (12·8)	<0·0001	679	35·7 (12·9)	36·3 (12·5)	0·070
SPPB ≤10¶	181 (38·9)	166 (46·0)	582 (58·8)	<0·0001	685	318 (46·4)	309 (45·1)	0·53
ISWT distance, m*	487·6 (274·7)	431·4 (242·3)	384·6 (249·4)	<0·0001	509	453·6 (262·8)	468·2 (267·8)	0·017
ISWT % predicted*	63·5 (30·7)	57·8 (28·1)	52·5 (28·7)	<0·0001	429	60·1 (29·4)	61·2 (28·7)	0·22
MoCA <23¶	66 (16·5)	39 (12·4)	147 (15·8)	0·26	623	89 (14·3)	62 (10·0)	0·0013
MoCA (adjusted) <23¶	57 (14·3)	36 (11·4)	125 (13·4)	0·52	623	72 (11·6)	55 (8·8)	0·034
FEV1 <80% predicted¶	43 (21·4)	43 (27·7)	131 (27·9)	0·20	287	67 (23·3)	63 (22·0)	0·64
FVC <80% predicted¶	40 (19·9)	33 (21·4)	155 (33·2)	0·00030	281	79 (28·1)	63 (22·4)	0·018
BNP ≥100 ng/L or Pr-NT-BNP ≥400 ng/L¶	27 (8·7)	24 (10·3)	35 (5·2)	0·014	335	30 (9·0)	29 (8·7)	1·0

HbA1C ≥6·0% (DCCT/NGSP)¶	6·1 (1·2)	6·2 (1·3)	6·2 (1·3)	0·60	399	140 (35·1)	130 (32·6)	0·21
eGFR < 60 ml/min/1·73 m²¶	49 (12·0)	35 (11·3)	101 (11·4)	0·94	564	73 (12·9)	79 (14·0)	0·45
CRP >5 mg/L¶	83 (20·9)	75 (23·9)	239 (27·1)	0·052	557	126 (22·6)	133 (23·9)	0·52
EQ5DL Utility Index †	0·88 (0·75 - 1·00)	0·77 (0·65 - 0·88)	0·69 (0·52 - 0·80)	<0·0001	585	0·74 (0·64 - 0·88)	0·75 (0·62 - 0·88)	0·95
EQ5D-5L VAS †	85·0 (72·2 - 91·2)	75·0 (60·0 - 85·0)	70·0 (50·0 - 80·0)	<0·0001	586	75·0 (60·0 - 90·0)	75·0 (60·0 - 90·0)	0·43
WG-SS-SCo †	0·0 (0·0 - 2·0)	2·0 (0·5 - 3·0)	3·0 (1·0 - 8·0)	<0·0001	548	2·0 (0·0 - 4·0)	2·0 (0·0 - 4·0)	0·73

¶ = Number (%) with positive response, *Mean [SD], †Median [IQR]. Percentages are calculated by category after exclusion of missing data for that variable. P values across patient perceived recovery were calculated using a chi-squared test when testing for differences between proportions, ANOVA F-test for normally distributed continuous data and Kruskal Wallis for non-normally distributed continuous data. For paired data, P values for McNemar's Chi-squared test with continuity correction for binary variables, P values for McNemar's Chi-squared test when there are more than two levels, P values for Wilcoxon signed-rank test for variables summarised as median [IQR] are presented and P values for paired t-test for variables summarised as mean [SD] are presented. VAS = Visual Analogue Scale 0-10, GAD7 = Generalized Anxiety Disorder 7-item scale. PHQ-9 = Patient Health Questionnaire-9. PCL-5 = Post Traumatic Stress Disorder Checklist. FACIT fatigue = Functional Assessment of Chronic Illness Therapy Fatigue Scale. BPI = Brief Pain Inventory, WG-SS-SCo = Washington Group Short Set of Functioning Severity Continuum, SPPB = short physical performance battery. ISWT = incremental shuttle walk test. CFS = Clinical Frailty Scale. MoCA = Montreal Cognitive Assessment. Adjusted MoCA = MoCA adjusted for education. FEV1 = Forced Expiratory Volume in the first second. FVC = forced vital capacity. TLCO = transfer capacity of the lung for carbon monoxide. KCO = carbon monoxide transfer coefficient. BNP = brain natriuretic peptide. Pro-NT-BNP = N-terminal BNP. HbA1C = glycated haemoglobin. DCCT/NGSP = Diabetes Control and Complications Trial/National Glycohemoglobin Standardization Program. eGFR = estimated glomerular filtration rate. CRP = C-reactive protein.

Table 3. A comparison of the change in patient reported outcome measures including mental health, physical function, cognitive impairment, and organ function between five-months and one-year sub-grouped by the four 'severity' clinical phenotypes

	1: Very severe				2: Severe				3: Moderate / Cognitive				4: Mild			
	N pairs	5-m	1-year	p	*N pairs*	5-m	1-year	p	*N pairs*	5-m	1-year	p	*N pairs*	5-m	1-year	p
Time to review, days †	*99*	174 (136 – 204)	393 (354 – 413)	-	*176*	176 (152 – 193)	386 (358 – 409)	-	*75*	174 (155 – 199)	378 (363 – 405)	-	*252*	180 (163 – 194)	388 (364 – 407)	-
PROMS																
Symptom count †	*80*	19.5 (16 – 25)	20 (14 – 26)	0·30	*141*	13 (8 – 17)	13 (8 – 18)	0·11	*61*	7.5 (5 – 11)	9 (5 – 12)	0·22	*188*	4 (1 – 7)	4 (1 – 7.5)	0·75
Anxiety (GAD7 ≥8)¶	*88*	93·2	70·5	0·00010	*152*	24·3	23·0	0·89	*70*	5·7	14·3	0·077	*225*	1·3	3·6	0·13
Depression (PHQ-9 ≥ 10)¶	*87*	96·6	80·5	0·0012	*152*	32·2	27·0	0·28	*70*	2·9	7·1	0·37	*225*	0·9	3·1	0·18
PTSD (PCL-5 ≥ 38)¶	*88*	65·9	47·7	0·0046	*152*	3·3	6·6	0·23	*68*	0·0	1·5	NA	*223*	0·0	0·4	NA
Dyspnoea-12 score*	*89*	18·0 (9·6)	15·2 (10·0)	0·0028	*167*	6·5 (5·8)	5·7 (5·9)	0·11	*67*	2·9 (3·7)	4 (5·4)	0·066	*238*	1·5 (2·4)	1·4 (2·5)	0·64
FACIT fatigue subscale score*	*88*	17·4 (8·9)	21·8 (12·0)	0·00010	*152*	30·6 (8·9)	33·5 (10·1)	<0·0001	*69*	42·0 (7·4)	39·1 (8·5)	0·0052	*225*	45·8 (4·8)	45·0 (6·6)	0·043
Physical performance																
SPPB ≤ 10¶	*86*	68·6	69·8	1·0	*164*	48·8	46·3	0·67	*70*	80·0	70·0	0·096	*239*	23·0	26·4	0·34
ISWT distance, m*	*65*	308 (225)	332 (233)	0·14	*122*	454 (262)	476 (286)	0·056	*45*	349 (200)	355 (155)	0·72	*172*	552 (254)	571 (268)	0·070
ISWT % predicted*	*58*	40·7 (24·7)	44·4 (27·2)	0·14	*105*	59·9 (27·3)	62·4 (30·5)	0·11	*39*	56·5 (30·8)	57·3 (26·1)	0·72	*138*	71·6 (27·9)	72·4 (26·3)	0·62
Cognitive impairment																
MOCA <23¶	*88*	22·7	15·9	0·15	*161*	5·6	5·6	1·0	*62*	59·7	32·3	0·00050	*232*	3·9	3·0	0·75
MOCA (corrected) <23¶	*88*	20·5	13·6	0·11	*161*	3·7	4·3	1·0	*62*	53·2	29·0	0·0023	*232*	1·7	2·6	0·68
Organ function																
FEV1 % <80% predicted¶	*28*	35·7	25·0	0·25	*61*	21·3	18·0	0·72	*33*	21·2	27·3	0·72	*86*	18·6	22·1	0·55
FVC % <80% predicted¶	*27*	51·9	37·0	0·13	*60*	25	18·3	0·22	*33*	36·4	27·3	0·45	*86*	22·1	23·3	1·0

CRP >5 mg/L¶	*73*	34·2	38·4	0·51	*118*	30·5	27·1	0·45	*52*	7·7	21·2	0·046	*165*	13·9	14·5	1·0

¶ = % of category with positive response, *Mean [SD], † median [IQR]. Missing not included in %. P values for McNemar's Chi-squared test with continuity correction for binary variables, P values for McNemar's Chi-squared test when there are more than two levels, P values for Wilcoxon signed-rank test for variables summarised as median [IQR] are presented and P values for paired t-test for variables summarised as mean [SD] are presented. Symptom count denominator n=XX. PROMS = Patients reported outcome measures. GAD7 = Generalized Anxiety Disorder 7-item scale. PHQ-9 = Patient Health Questionnaire-9. PCL-5 = Post Traumatic Stress Disorder Checklist. FACIT fatigue = Functional Assessment of Chronic Illness Therapy Fatigue Scale. BPI = Brief Pain Inventory. SPPB = short physical performance battery. ISWT = incremental shuttle walk test. CFS = Clinical Frailty Scale. MoCA = Montreal Cognitive Assessment. Adjusted MoCA = MoCA adjusted for education. FEV1 = Forced Expiratory Volume in the first second. FVC = forced vital capacity. TLCO = transfer capacity of the lung for carbon monoxide. KCO = carbon monoxide transfer coefficient. BNP = brain natriuretic peptide. NT-BNP = N-terminal BNP. HbA1C = glycated haemoglobin. DCCT/NGSP = Diabetes Control and Complications Trial/National Glycohemoglobin Standardization Program. eGFR = estimated glomerular filtration rate. CRP = C-reactive protein.

Figure 1

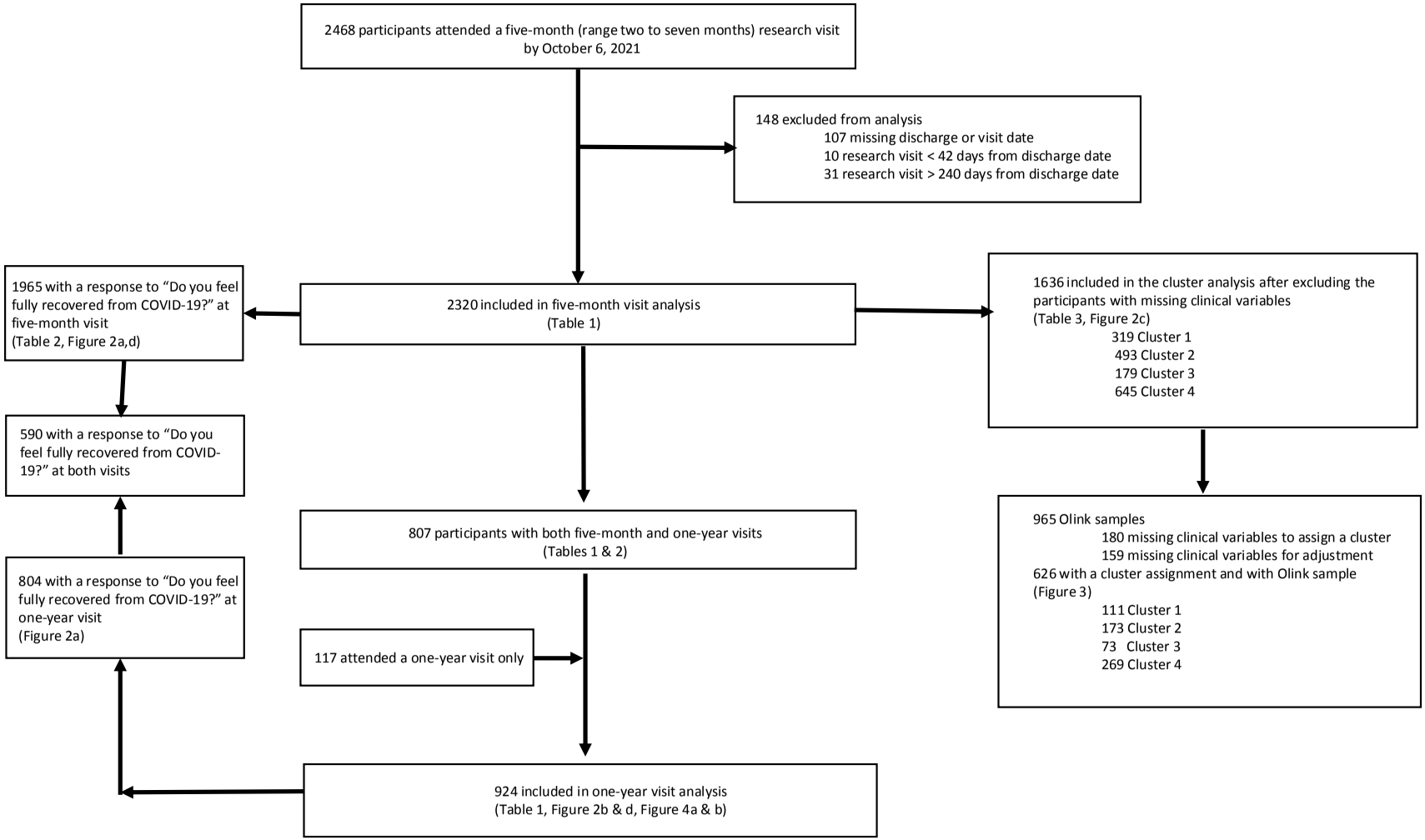


Figure 1: CONSORT diagram

Figure 2

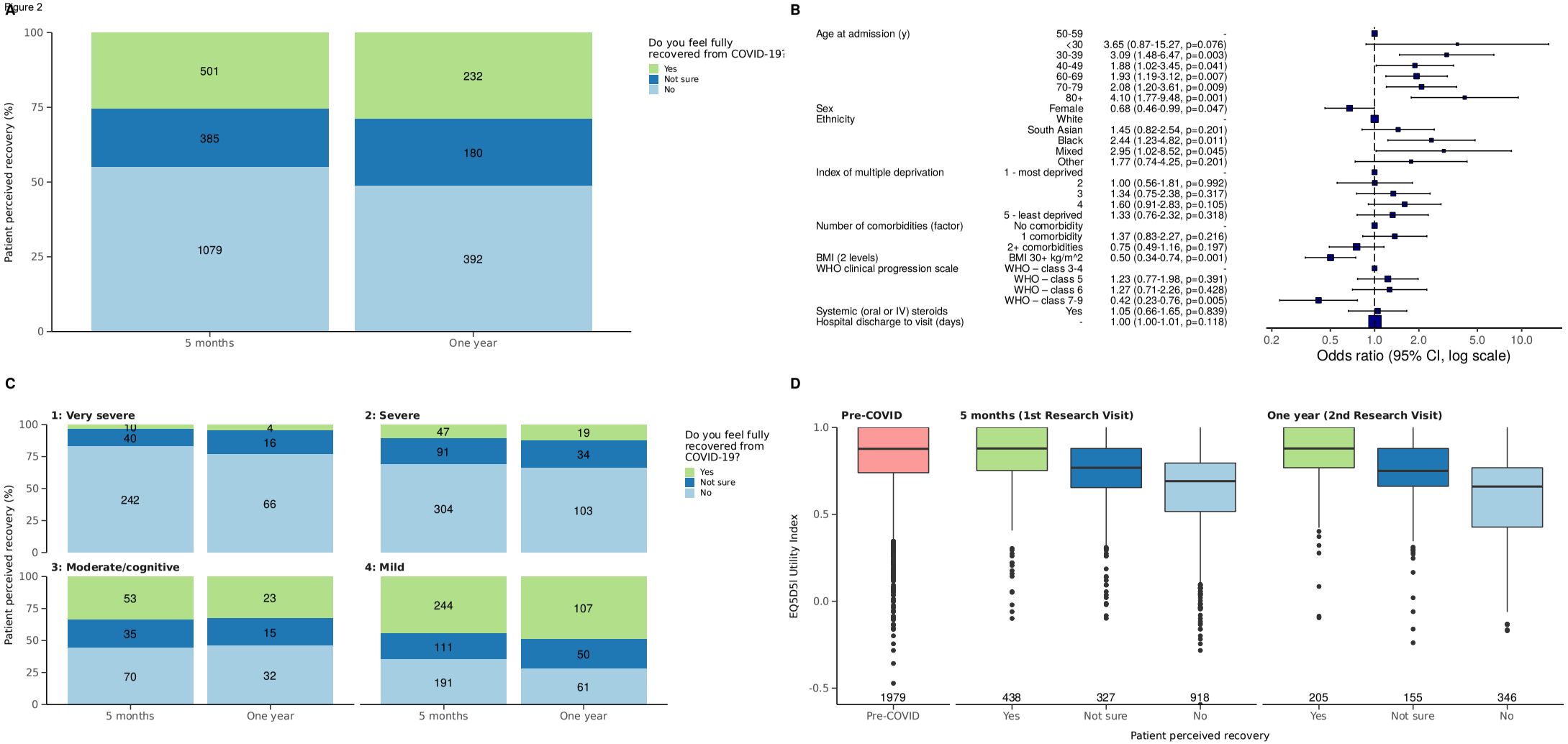

A
Patient perceived recovery (%)
100
75
50
25
0
501
385
1079
232
180
392
5 months
One year
Do you feel fully recovered from COVID-19?
Yes
Not sure
No
B
Age at admission (y)
50-59 -
<30 3.65 (0.87-15.27, p=0.076)
30-39 3.09 (1.48-6.47, p=0.003)
40-49 1.88 (1.02-3.45, p=0.041)
60-69 1.93 (1.19-3.12, p=0.007)
70-79 2.08 (1.20-3.61, p=0.009)
80+ 4.10 (1.77-9.48, p=0.001)
Sex
Female 0.68 (0.46-0.99, p=0.047)
Ethnicity
White -
South Asian 1.45 (0.82-2.54, p=0.201)
Black 2.44 (1.23-4.82, p=0.011)
Mixed 2.95 (1.02-8.52, p=0.045)
Other 1.77 (0.74-4.25, p=0.201)
Index of multiple deprivation
1 - most deprived -
2 1.00 (0.56-1.81, p=0.992)
3 1.34 (0.75-2.38, p=0.317)
4 1.60 (0.91-2.83, p=0.105)
5 - least deprived 1.33 (0.76-2.32, p=0.318)
Number of comorbidities (factor)
No comorbidity -
1 comorbidity 1.37 (0.83-2.27, p=0.216)
2+ comorbidities 0.75 (0.49-1.16, p=0.197)
BMI (2 levels)
BMI 30+ kg/m^2 0.50 (0.34-0.74, p=0.001)
WHO clinical progression scale
WHO – class 3-4 -
WHO – class 5 1.23 (0.77-1.98, p=0.391)
WHO – class 6 1.27 (0.71-2.26, p=0.428)
WHO – class 7-9 0.42 (0.23-0.76, p=0.005)
Systemic (oral or IV) steroids
Yes 1.05 (0.66-1.65, p=0.839)
Hospital discharge to visit (days)
- 1.00 (1.00-1.01, p=0.118)
0.2
0.5
1.0
2.0
5.0
10.0
Odds ratio (95% CI, log scale)
C
1: Very severe
2: Severe
3: Moderate/cognitive
4: Mild
Patient perceived recovery (%)
10
40
242
4
16
66
47
91
304
19
34
103
53
35
70
23
15
32
244
111
191
107
50
61
5 months
One year
Do you feel fully recovered from COVID-19?
Yes
Not sure
No
D
Pre-COVID
5 months (1st Research Visit)
One year (2nd Research Visit)
EQ5D5l Utility Index
1.0
0.5
0.0
-0.5
1979
438
327
918
205
155
346
Pre-COVID
Yes
Not sure
No
Yes
Not sure
No
Patient perceived recovery

Figure 13: Very severe mental/physical impairment versus Mild

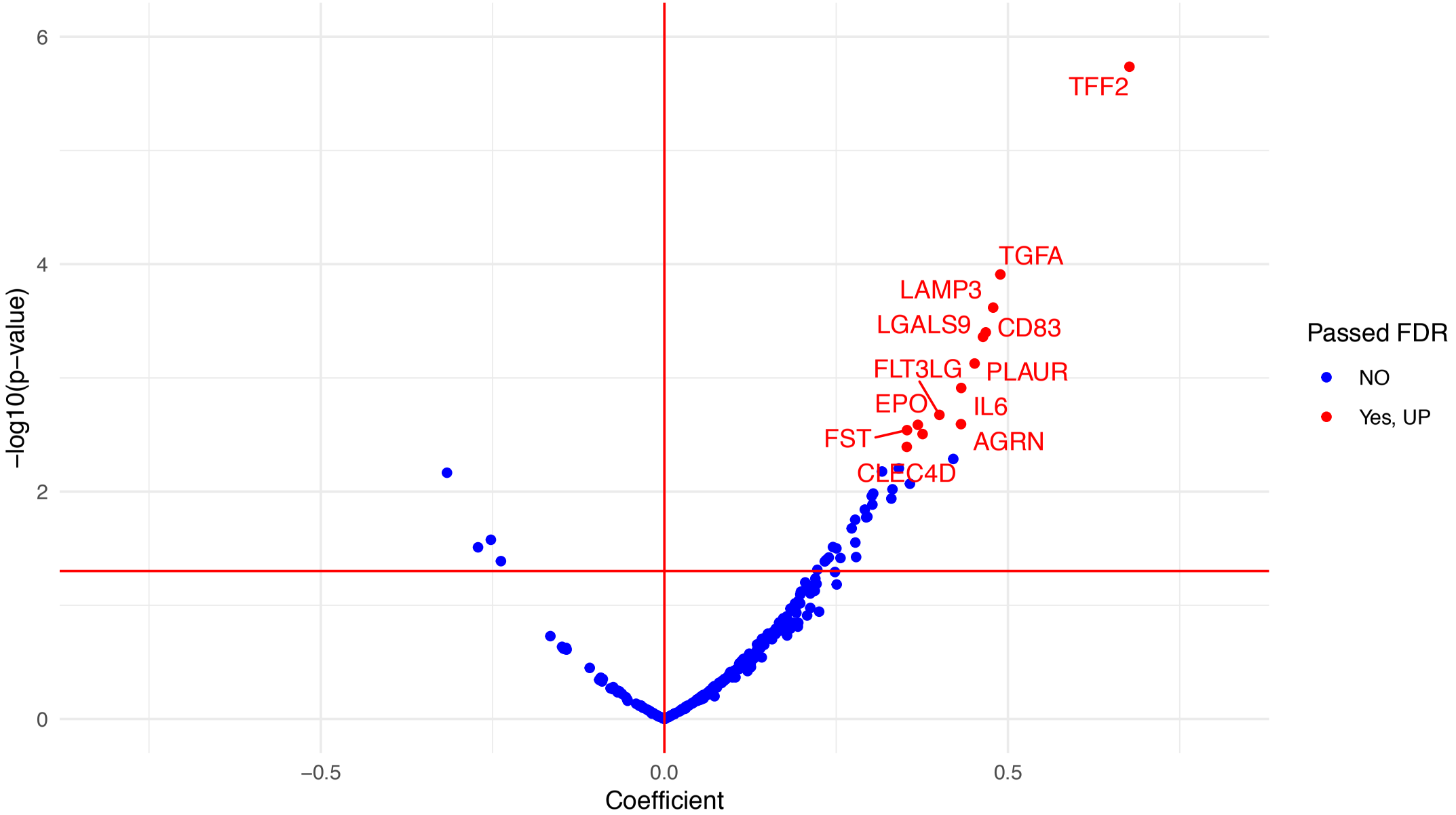


Figure 23b

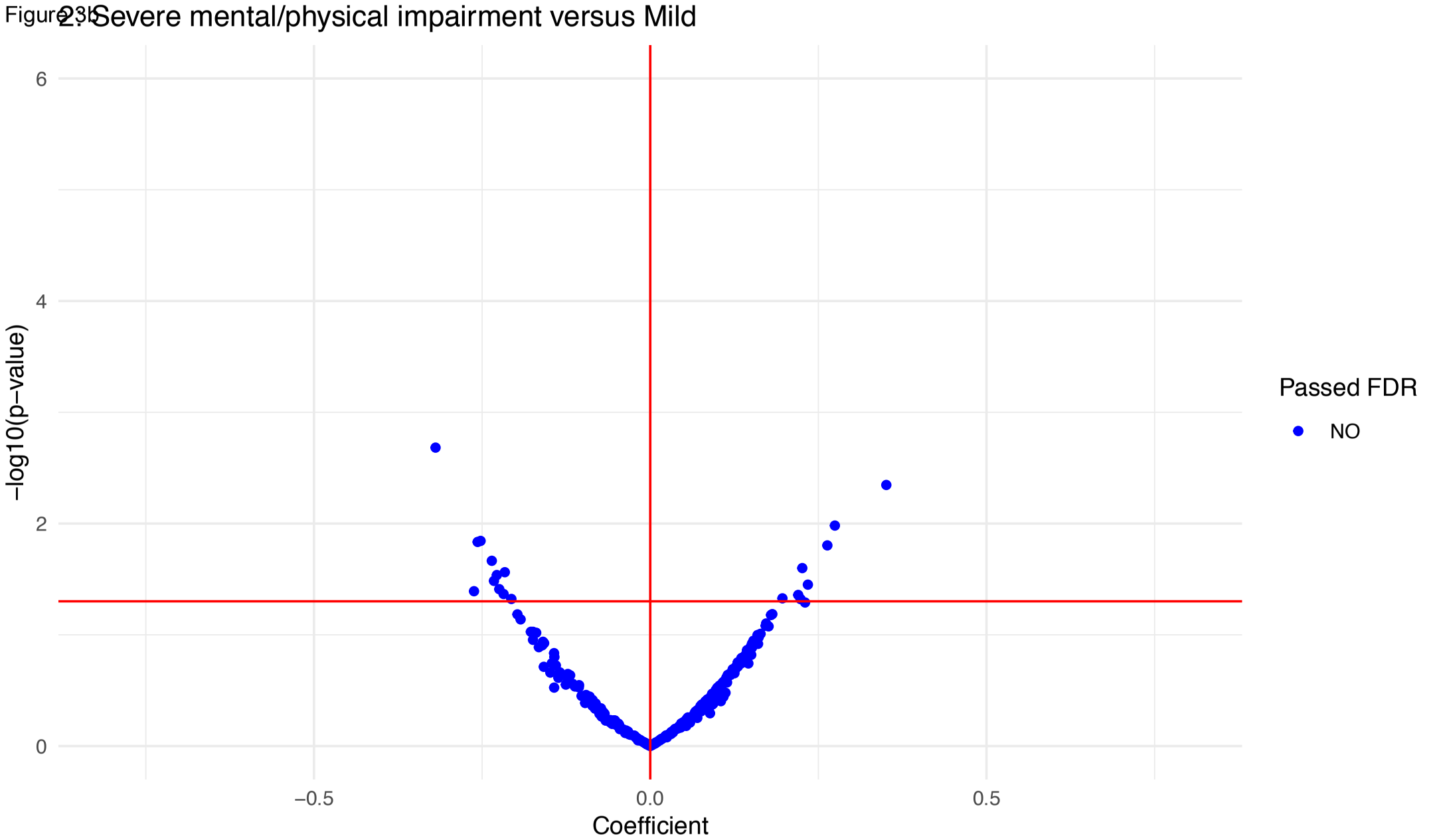

2. Severe mental/physical impairment versus Mild
6
4
2
0
−log10(p−value)
−0.5
0.0
0.5
Coefficient
Passed FDR
NO

Figure 33:

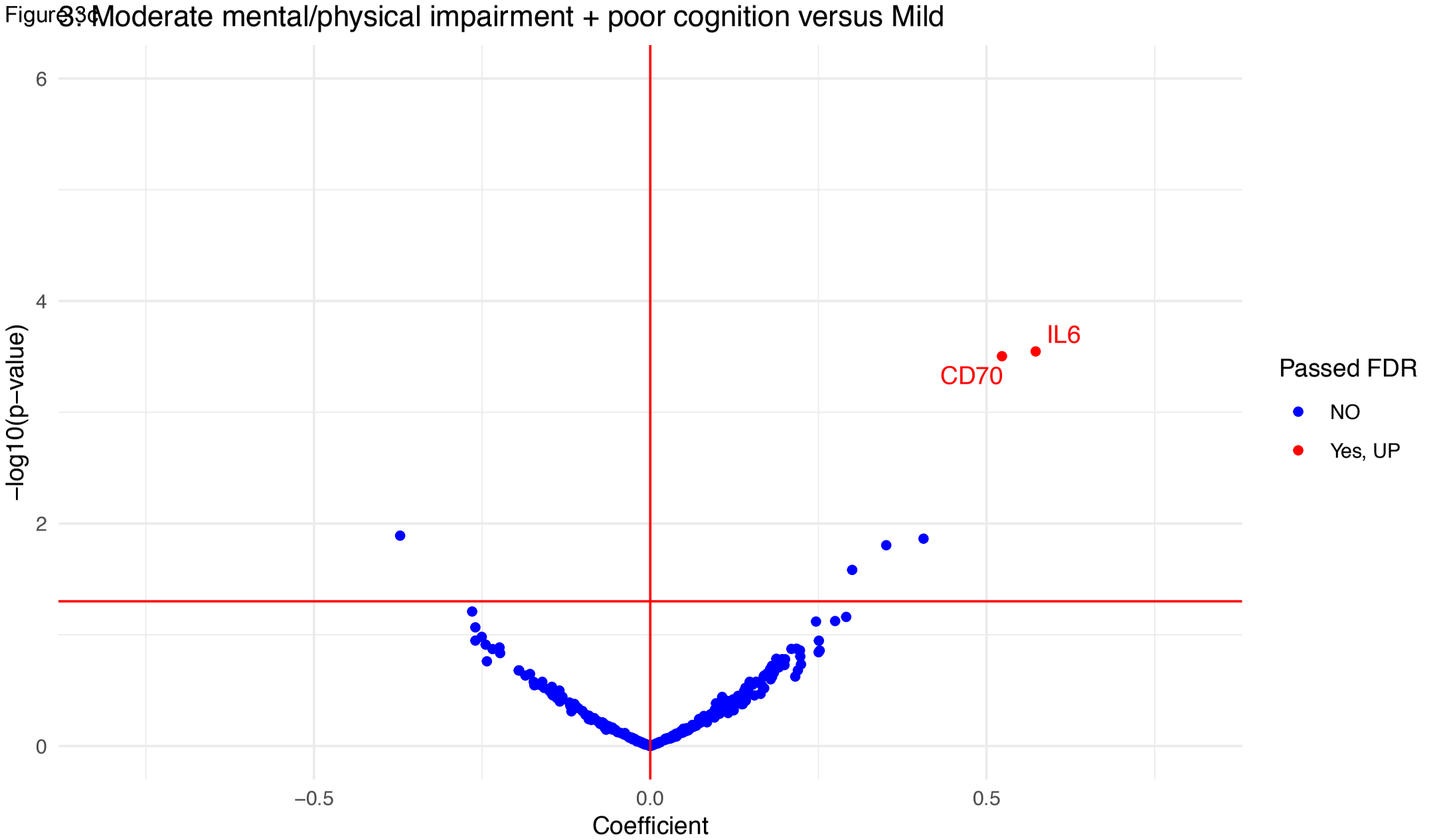

3: Moderate mental/physical impairment + poor cognition versus Mild
6
4
2
0
−log10(p−value)
IL6
CD70
−0.5
0.0
0.5
Coefficient
Passed FDR
NO
Yes, UP

Figure 4a

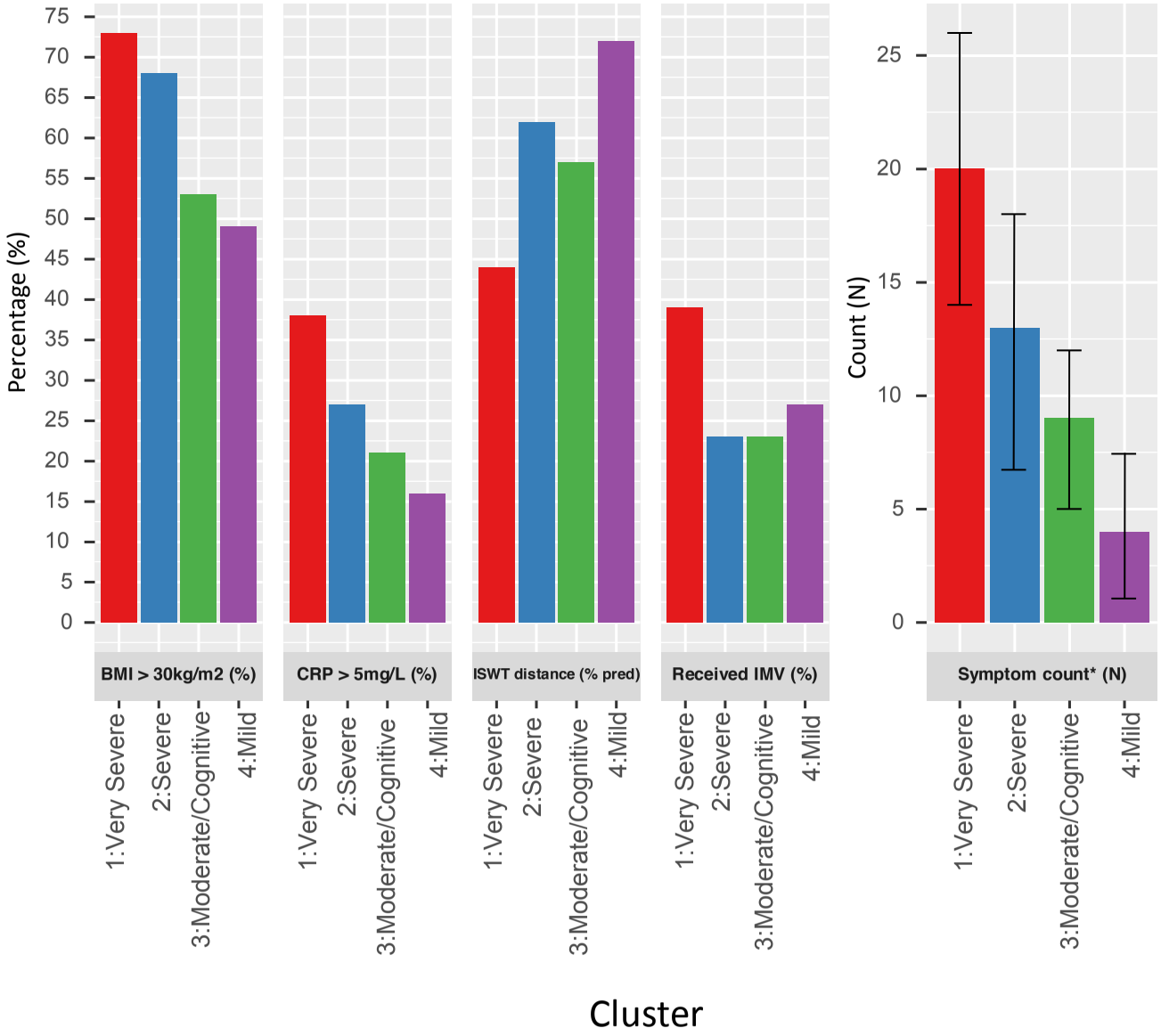
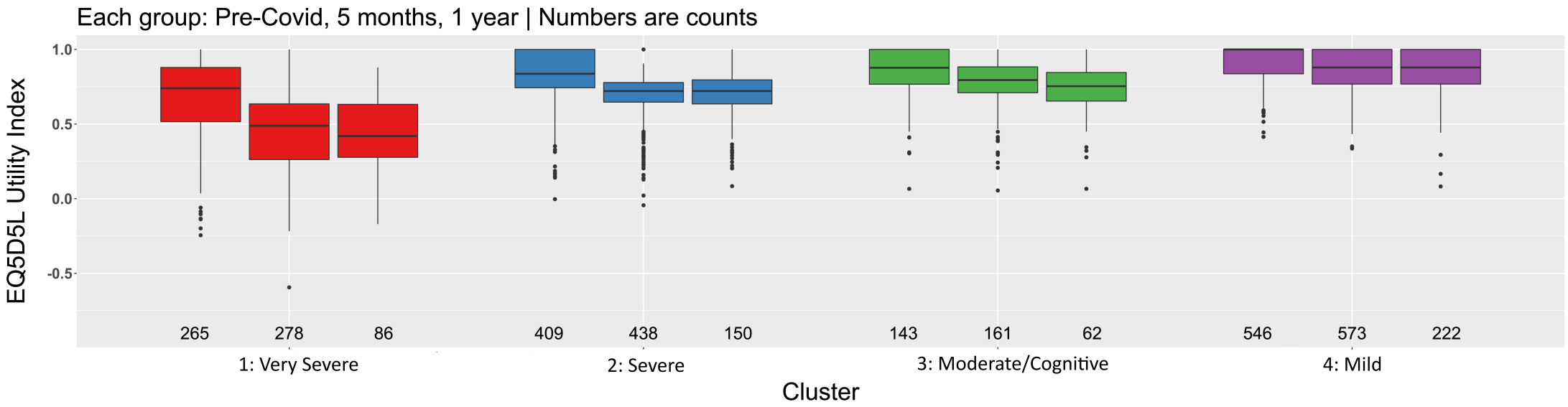

Percentage (%)
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
BMI > 30kg/m2 (%)
CRP > 5mg/L (%)
ISWT distance (% pred)
Received IMV (%)
Count (N)
25
20
15
10
5
0
Symptom count* (N)
1:Very Severe
2:Severe
3:Moderate/Cognitive
4:Mild
Cluster

Figure 4b


Each group: Pre-Covid, 5 months, 1 year | Numbers are counts
EQ5D5L Utility Index
1.0
0.5
0.0
-0.5
265
278
86
409
438
150
143
161
62
546
573
222
1: Very Severe
2: Severe
3: Moderate/Cognitive
4: Mild
Cluster