Large scale phenotyping of long COVID inflammation reveals mechanistic subtypes of disease after COVID-19 hospitalisation

Peter Openshaw

p.openshaw@imperial.ac.uk

Imperial College London https://orcid.org/0000-0002-7220-2555

Felicity Liew

Imperial College

Claudia Efstathiou

NHLI, Imperial College London https://orcid.org/0000-0001-6125-8126

Sara Fontanella

National Heart and Lung Institute, Imperial College London

Matthew Richardson

University of Leicester

Ruth Saunders

University of Leicester

Dawid Swieboda

NHLI, Imperial College London

Jasmin Sidhu

NHLI, Imperial College London

Stephanie Ascough

Imperial College London

Shona Moore

University of Liverpool https://orcid.org/0000-0001-8610-2806

Noura Mohamed

NHLI, Imperial College London

Jose Nunag

NHLI, Imperial College London

Clara King

NHLI, Imperial College London

Olivia Leavy

University of Leicester

Omer Elneima

University of Leicester

Hamish McAuley

University of Leicester

Aarti Shikotra

University of Leicester

Amisha Singapuri

University of Leicester https://orcid.org/0009-0002-4711-7516

Marco Sereno

University of Leicester https://orcid.org/0000-0003-4573-9303
Victoria Harris
University of Leicester
Linzy Houchen-Wolloff
Leicester https://orcid.org/0000-0003-4940-8835
Neil Greening
University of Leicester https://orcid.org/0000-0003-0453-7529
Nazir Lone
University of Edinburgh https://orcid.org/0000-0003-2707-2779
Matthew Thorpe
Usher Institute, University of Edinburgh, Edinburgh, UK
A.A. Thompson
University of Sheffield https://orcid.org/0000-0002-0717-4551
Sarah Rowland-Jones
University of Sheffield
Annemarie Docherty
University of Edinburgh https://orcid.org/0000-0001-8277-420X
James Chalmers
University of Dundee, Ninewells Hospital and Medical School, Dundee, Scotland
Ling-Pei Ho
University of Oxford https://orcid.org/0000-0001-8319-301X
Alex Horsley
University of Manchester https://orcid.org/0000-0003-1828-0058
Betty Raman
University of Oxford
Krisnah Poinasamy
Asthma and Lung UK
Michael Marks
London School of Hygiene & Tropical Medicine
Onn Min Kon
NHLI, Imperial College London
Luke Howard
National Pulmonary Hypertension Service, Imperial College Healthcare Trust NHS https://orcid.org/0000-0003-2822-210X
Jennifer Quint
National Heart and Lung Institute, Imperial College London, London, United Kingdom
Thushan de Silva
The University of Sheffield, Sheffield Teaching Hospitals NHS Foundation Trust https://orcid.org/0000-0002-6498-9212
Antonia Ho
MRC-University of Glasgow Centre for Virus Research
Christopher Chiu
Imperial College London https://orcid.org/0000-0003-0914-920X
Ewen Harrison
University of Edinburgh https://orcid.org/0000-0002-5018-3066
William Greenhalf
Royal Liverpool University Hospital
J Baillie
Roslin Institute, University of Edinburgh https://orcid.org/0000-0001-5258-793X
Malcolm Semple

NIHR Health Protection Research Unit in Emerging and Zoonotic Infections, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool

Rachael Evans
University of Leicester https://orcid.org/0000-0002-1667-868X

Louise Wain
University of Leicester

Christopher Christopher
University of Leicester

Lance Turtle
Institute of Infection, Veterinary and Ecological Sciences, Faculty of Health and Life Sciences, University of Liverpool, Liverpool, UK

Daniel Wootton
NIHR Health Protection Research Unit in Emerging and Zoonotic Infections, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool

Ryan Thwaites
Imperial College London https://orcid.org/0000-0003-3052-2793

Letter

Keywords: COVID-19, SARS-CoV-2, inflammation, post-viral sequelae, innate immunity, complement, long COVID, PASC, proteomics

Posted Date: December 4th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3427282/v1

Additional Declarations: Yes there is potential Competing Interest. PJMO reports grants from the EU Innovative Medicines Initiative (IMI) 2 Joint Undertaking during the submitted work; grants from UK Medical Research Council, GlaxoSmithKline, Wellcome Trust, EU-IMI, UK, National Institute for Health Research, and UK Research and Innovation-Department for Business, Energy and Industrial Strategy; and personal fees from Pfizer, Janssen, and Seqirus, outside the submitted work.

Version of Record: A version of this preprint was published at Nature Immunology on April 8th, 2024. See the published version at https://doi.org/10.1038/s41590-024-01778-0.

Abstract

One in ten SARS-CoV-2 infections result in prolonged symptoms termed long COVID, yet disease phenotypes and mechanisms are poorly understood. We studied the blood proteome of 719 previously hospitalised adults with long COVID grouped by symptoms. Elevated markers of myeloid inflammation and complement activation were associated with long COVID; elevated IL1R2, MATN2 and COLEC12 were associated with cardiorespiratory symptoms, fatigue, and anxiety/depression, while MATN2 and DPP10 were elevated in gastrointestinal (GI) symptoms, and C1QA in cognitive impairment. Proteins suggestive of neurodegeneration were elevated in cognitive impairment, whilst SCG3 (indicative of brain-gut axis disturbance) was specific to GI symptoms. Nasal inflammation was apparent after COVID-19 but did not associate with symptoms. Although SARS-CoV-2 specific IgG was elevated with some long COVID symptoms, virus was not detected from sputum. Thus, systemic inflammation is evident in long COVID and could be targeted in therapeutic trials tailored to pathophysiological differences between symptom groups.

Main

One in ten SARS-CoV-2 infections results in long COVID or Post-acute sequelae of COVID-19 (PASC) affecting 65 million worlwide.[1,2] SARS-CoV-2 continues to circulate and long COVID remains common, even after mild acute infection with recent variants.[3,4] It is likely long COVID will continue to cause substantial long-term ill health, requiring targeted management based on an understanding of how disease phenotypes relate to underlying mechanisms.

There are reports of persistent inflammation in adults with long COVID,[5–9] but studies have been limited by size, timing of samples or breadth of immune mediators measured, leading to inconsistent or absent associations with symptoms.[10] Markers of oxidative stress, metabolic disturbance, vasculoproliferative processes and IFN-, NF-kB- or monocyte-related inflammation have been suggested.[5,7,11–13] The PHOSP-COVID study recently reported the plasma proteome of 626 adults with long COVID (identified through clustering and utilising measures of breathlessness, fatigue, mental health, cognitive impairment, and physical performance). Elevated IL-6 and markers of mucosal inflammation were observed in those with severe long COVID compared to individuals with milder symptoms.[14,15] Long COVID is a heterogenous condition, that is not well defined, and it remains unclear if there are inflammatory changes specific to symptom type, when compared to recovered controls.

Many long COVID symptoms have been described, most commonly breathlessness, fatigue, memory impairment and gastrointestinal (GI) disturbance.[16–18] There are reports of SLE-like autoantibodies in individuals with GI symptoms, whilst Epstein-Barr Virus (EBV) reactivation has been associated with fatigue and neurological symptoms,[6,19–21] suggesting that distinct mechanisms might cause different symptom patterns and that these might be revealed by analysis of inflammatory markers. Confirming whether common inflammatory pathways underlie all cases of long COVID or if mechanisms differ according to clinical presentation, is essential for developing effective therapeutic approaches and has been highlighted as a top research priority by patients, clinicians and scientists.[22] In this prospective multicentre study, we measured 368 plasma proteins in 719 adults previously hospitalised for COVID-19, to understand inflammatory processes underlying common long COVID symptoms.

Results

We studied 719 adults from the PHOSP-COVID study, including the 626 patients previously reported.[15] Individuals had been hospitalised for COVID-19, 6 months prior (median 6.0; IQR 4.5–6.7 months; Range 1.4–8.3), confirmed clinically or by PCR (n=621). We included patients reporting symptoms from 4 weeks after acute COVID-19, per the National Institute for Health and Care Excellence (NICE) and Centers for Disease Control and Prevention (CDC) definitions of long COVID (Fig. 1A).[23,24] Analysing cross-sectional clinical data, 250/719 (35%) felt fully recovered (“Recovered”) and the remaining 469 (65%) reported symptoms consistent with long COVID (Fig. 1B; Table 1).

Using a multivariate penalised logistic regression model (PLR) to explore the associations between clinical covariates, immune mediators and symptoms, we found women were more likely to experience all symptoms, and this effect was largest for GI

(Odds Ratio; OR=1.13) and cardiorespiratory symptoms (OR=1.17; Fig. 1C–G). Confidence intervals are not appropriately derived from PLR analysis and are not reported, however repeated cross-validation was used to estimate uncertainty associated with PLR outputs (Methods and Supplementary). Pre-existing conditions that might predispose to symptom outcomes (e.g., chronic lung disease in the case of cardiorespiratory symptoms; Supplementary Table 1) were associated with all symptoms, except GI. Age and acute disease severity were not associated with any symptom. We did not include ethnicity as a covariate because it is not an independent risk factor in this cohort.[16]

Myeloid inflammation and complement activation are common to all long COVID symptoms

To study the association of peripheral inflammation with symptoms, 368 immune mediators were measured from plasma and included as covariates. Mediators suggestive of myeloid inflammation were associated with all symptoms (Fig. 1C–G). Elevated IL1R2 and/or Matrilin-2 (MATN2) were consistently associated with the highest odds of all symptoms, except cognitive impairment where the effect was smaller (cardiorespiratory IL1R2 OR=1.16; fatigue IL1R2 OR=1.53; anxiety/depression IL1R2 OR=1.13; GI MATN2 OR=1.08; cognitive MATN2 OR=1.03). IL1R2 is expressed by monocytes and macrophages, modulating IL-1 inflammation.[25] MATN2 is an extracellular matrix (ECM) protein which promotes inflammation by activating toll-like receptors and enhancing monocyte infiltration into tissues.[26,27] CSF3 (G-CSF, which promotes neutrophilic inflammation), was elevated in fatigue (OR=1.12), GI symptoms (OR=1.05) and anxiety/depression (OR=1.05; Fig. 1D–F).[28] Increased levels of IL-6 were associated with cardiorespiratory symptoms (OR=1.06) and fatigue (OR=1.09).

Elevated Collectin-12 (COLEC12) was also associated with cardiorespiratory symptoms (OR=1.12), anxiety/depression (OR=1.06) and fatigue (OR=1.21; Fig. 1C–E). COLEC12 can initiate inflammation in tissues by activating the alternative complement pathway.[29,30] Whilst COLEC12 was not associated with GI symptoms and only weakly associated with cognitive impairment (OR=1.02), C1QA was associated with these symptoms (Fig. 1F&G). C1QA is a component of the complement system, indicating activation via the classical pathway.[31] Notably, C1QA was associated with the second highest odds of cognitive impairment (OR=1.04), and has been implicated in the pathogenesis of chronic neuroinflammation in Alzheimer's disease.[32] Although subtle differences were observed between symptom groups, our findings demonstrate myeloid inflammation and complement activation in all long COVID phenotypes.

We used the CDC and NICE definition for long COVID (>4 weeks after acute COVID-19) in our analyses.[23,24] However, the World Health Organisation (WHO) defines long COVID as symptoms occurring 3 months post-infection.[33] We therefore repeated our analysis using samples and clinical data collected after 3 months (Median 6.1 months; IQR 5.1-6.8; Range 3.0-8.3; n=659; recovered=233[35%]). Inflammatory associations with long COVID symptoms were consistent with our original analysis, indicating that the profiles identified in our cohort were representative of long COVID after hospitalisation using three commonly used definitions (Extended Data Fig. 1A-G).

To further validate the findings from PLR analysis, we examined the distribution of data, prioritising proteins that were associated with the highest odds of each symptom (Fig. 1H–L and Extended Data Fig. 2). Each protein was significantly elevated in the symptom group compared to recovered, confirming the patterns identified by PLR. Unadjusted PLR models and alternative regression approaches (Partial Least Squares; PLS) were also used to confirm the validity of our findings (Supplementary Table 2 and Extended Data Fig. 3,4). Results from these approaches confirmed the relationship between female sex and comorbidities on outcome, as well as the association between myeloid inflammation, complement and symptoms. Notably the standard errors of PLS estimates were wide, consistent with the literature, reporting PLR as the optimal method to analyse multiple mediators which may correlate due to their combined effects.[34] Since we aimed to understand how inflammatory proteins work together to mediate symptoms, we prioritised PLR results to draw conclusions.

Biomarker discovery was not our goal and the marked overlap in mediator levels when viewed unidimensionally, indicates these markers are not useful on an individual basis for diagnosis (Fig. 1H–M). Importantly, we did not find differences in C-reactive protein (CRP) levels between groups, measured contemporaneously by hospital laboratories (Table 1). Fibrinogen levels during acute COVID-19 have recently been associated with cognitive deficits post-COVID. [35,36] We similarly found that elevated fibrinogen was evident in long COVID (p=0.0077), suggesting that elevated fibrinogen in both the acute and post-acute phase

associates with long COVID symptoms (Extended Data Fig 1H). Given the interaction between complement activation and thrombosis, elevated fibrinogen supports our observation of complement pathway activation.[37]

GI symptoms and Cognitive impairment are associated with different patterns of inflammation

Whilst the protein signatures of individuals with cardiorespiratory symptoms, fatigue and anxiety/depression (the most common combination, n=88) were similar, specific proteins were raised in those with GI symptoms and cognitive impairment (Fig. 1F,G). Elevated Dipeptidyl peptidase 10 (DDP10) and Secretogranin 3 (SCG3) was observed in the GI group (DPP10 OR=1.07; SCG3 OR=1.06). DDP10 can modulate tissue inflammation, and increased *DPP10* expression is associated with Ulcerative Colitis, suggesting that GI symptoms may result from enteric, as well as systemic, inflammation.[38,39] Elevated SCG3 suggests disturbance of the brain-gut axis, as observed in patients with irritable bowel syndrome.[40]

Cognitive impairment was associated with elevated Neurofascin (NFASC; OR=1.05), Spondin-1 (SPON-1; OR=1.03) and Iduronate sulfatase (IDS; OR=1.04)(Fig. 1G,L). NFASC and SPON-1 regulate neural growth,[41,42] whilst IDS is an ECM enzyme supporting tissue turnover and enabling leucocyte infiltration into tissues.[43,44] The combination of these proteins with elevated C1QA, suggest neuroinflammation and alterations in nerve tissue repair (i.e., neurodegeneration). Taken together our findings indicate that complement activation and myeloid inflammation is common to all long COVID cases, but subtle differences in those with GI and Cognitive symptoms may have mechanistic significance.

Given our observations of elevated C1QA and the recent identification of acute fibrinogen as a biomarker of post-COVID cognitive impairment,[36] we analysed fibrinogen specifically in this group. We found that median fibrinogen levels were higher at 6 months in those with cognitive impairment (p=0.07), though this difference was not significant (Extended Data Fig 1I).

To explore the relationship between inflammatory mediators associated with different long COVID symptoms, we performed a network analysis of those mediators highlighted by PLR within each symptom group. COLEC12 and MATN2 showed high centrality compared to other mediators in the Cardiorespiratory, Fatigue and Anxiety/Depression groups (Fig. 2A–C & Extended Data Fig. 5A–C). Both mediators correlated with pro-inflammatory proteins (e.g., IL1R2, IL-12B [also known as IL-12/23p40], IL-6, CD276, CD4, DPP10) and markers of endothelial and mucosal inflammation (e.g., TGFA, TFF2, ISM1, ANGPTL2), suggesting roles in tissue-specific long COVID inflammation. Similarly, MATN2 and the pro-inflammatory protein TNFRSF11B were central to inflammation in the GI group (Extended Data Fig. 5D). However, SCG3 correlated less closely with mediators in this group, suggesting that alterations in the brain-gut axis may contribute separately to symptoms (Fig. 2D). SPON-1 was the most central mediator in those with cognitive impairment, further highlighting the possibility that neurodegenerative processes may occur in these individuals (Fig. 2E & Extended Data Fig. 5E). Taken together, these findings support the central role of complement and myeloid inflammation in long COVID but suggest additional processes may contribute towards GI symptoms and cognitive impairment.

Elevated sCD58 is associated with recovery

Elevated sCD58 was associated with lower odds of all long COVID symptoms and this was most pronounced for cardiorespiratory symptoms (OR=0.79; Fig. 1C,M). sCD58 is an immunoregulatory factor, known to suppress IL-1 and IL-6 dependent interactions between CD2+ monocytes and CD58+ (lymphocyte-function antigen 3) T/NK cells.[45,46] Since we observed markers of monocytic inflammation in all symptom groups and elevated IL-6 in those with fatigue and cardiorespiratory symptoms, the association of sCD58 and recovery supports the central role of myeloid inflammation in long COVID.

Elevated markers of tissue repair, including Delta/notch-like EGF repeat (DNER OR=0.82) were also associated with reduced risk of all symptoms (Fig. 1C–G). Notably, elevated IDS was associated with recovery compared to all symptom groups, except cognitive impairment where the inverse was true. IDS maintains tissues by preventing accumulation of ECM proteoglycans and facilitating leucocyte entry.[43,44] IDS may have divergent functions in different tissue environments, for example supporting lung tissue repair to prevent respiratory symptoms, whilst promoting neuroinflammation and thus cognitive impairment. Our data

suggests immunosuppressive factors and a robust tissue repair response may prevent symptoms after COVID-19, supporting the use of anti-inflammatory agents in therapeutic trials.[47]

Women who experience long COVID have higher inflammatory markers

We next sought to understand inflammatory responses in women, who were more likely to experience long COVID, in keeping with previous studies (Fig. 1C–G; Table 1).[16,18] Since oestrogen can influence immunological responses,[48] we compared protein levels between men and women younger and older than 50 years to discriminate between pre- and post-menopausal women (Fig. 3A–E). IL1R2 and MATN2 were significantly higher in women >50 years, with cardiorespiratory symptoms (IL1R2 p=0.0002; MATN2 p<0.0001), fatigue (IL1R2 p=0.0003; MATN2 p=0.012) and anxiety/depression (IL1R2 p=0.0003; MATN2 p=0.012). Oestrogen-dependent differences would be expected to be most pronounced in pre-menopausal women,[49] but this was not observed. Women have been reported to have stronger innate immune responses to infection[48,50] and are at greater risk of autoimmunity,[48] possibly explaining our findings.

Examining proteins associated with GI symptoms, there were no significant differences seen between men and women (Extended Data Fig. 6). In the cognitive impairment group IDS was significantly higher in pre-menopausal women (p=0.02), though this effect was lost in the post-menopausal group. *IDS* is X-linked, which may partially explain these differences.[51] Overall, our analysis suggests non-hormonal differences in immune responses explain the increased likelihood of women to experience long COVID. These findings require confirmation in adequately powered studies but have potential clinical implications, suggesting anti-inflammatory therapies might be most beneficial for women.

Systemic inflammation in long COVID is not related to the upper respiratory tract

We next sought to understand mechanisms driving long COVID inflammation, focussing on the cardiorespiratory group as the most common phenotype. Given the correlations observed between MATN2 and markers of mucosal inflammation in individuals with cardiorespiratory symptoms (Fig. 2A), we considered local inflammation in the respiratory tract as a possible cause. We analysed nasosorption samples from 88 adults within our cohort and 25 healthy controls (Supplementary Table 3). Several inflammatory markers were elevated in the upper respiratory tract post-COVID, including IL-1α (Fig. 4A). However, there was no difference between those recovered (n=31) and those not (n=33) (Fig. 4B). In adults with only cardiorespiratory symptoms (n=29), inflammatory mediators elevated in plasma were not elevated in the upper respiratory tract (Extended Data Fig. 7A–F). Furthermore, there was no correlation between mediator levels at different sites (Extended Data Fig. 7G–L). This exploratory analysis suggests that upper respiratory tract inflammation is not associated with cardiorespiratory symptoms.

Long COVID is associated with stronger antibody responses but not persistent sputum antigen

We next considered that SARS-CoV-2 persistence in lung tissue, might explain the inflammatory profiles observed in those with cardiorespiratory symptoms. We performed an exploratory analysis of SARS-CoV-2 antigens (S and N) in sputum from a subgroup of 23 adults with cardiorespiratory symptoms at 6 months. Sputum from 17 recovered adults and pre-pandemic bronchoalveolar lavage fluid were analysed as controls (Supplementary Table 3). Although low concentrations of N antigen were detected in 4 samples, there was no difference between those with symptoms and those recovered (Fig. 4C). S antigen was undetectable in all sputum samples.

Our findings do not exclude persistence, which is most likely evident from tissue samples.[52,53] We therefore examined SARS-CoV-2 specific antibody levels in a subgroup of unvaccinated individuals, which might respond to viral reservoirs. Consistent with previous reports, we found stronger SARS-CoV-2-specific IgG responses in individuals with persistent symptoms (Fig. 4D–H).[52] Both anti-S and -N IgG responses were higher in the Cardiorespiratory (S p=0.0040, Fig. 4D; N p =0.023, Fig. 4E) and Fatigue groups (S p=0.0030, Fig. 4F; N p=0.010, Fig. 4G), relative to Recovered. Anti-S (p=0.0098, Fig. 4H) but not -N (p=0.054, Fig. 4I) IgG was elevated in the Anxiety/Depression group. We did not have sufficient data to assess responses in Cognitive impairment and GI groups.

Overall, we demonstrate complement and myeloid associated inflammation in long COVID alongside elevated antibody titres, providing insights into disease mechanisms and aetiology.

Discussion

In this study of 719 adults who survived hospitalisation for COVID-19, we demonstrate myeloid inflammation and complement activation in patients experiencing cardiorespiratory symptoms, fatigue, anxiety/depression, cognitive impairment and GI symptoms. Our findings build on results of smaller studies[12,35,54] and are consistent with a recent GWAS study identifying an independent association between long COVID and *FOXP4*, which modulates neutrophilic inflammation and immune cell function.[55,56] We identified tissue-specific inflammatory elements, indicating that ongoing myeloid disturbance may be occurring in different tissues, resulting in distinct symptoms. This could reflect persistent SARS-CoV-2 antigen within these tissues, or a failure to resolve inflammation and immune activation post-infection.

We highlight complement activation in long COVID, which has been suggested by one smaller proteomic study of mostly non-hospitalised COVID cases.[35] This is significant since components of the complement system are known to have a short half-life,[57] suggesting symptoms result from active inflammation and not passively from tissue damage incurred by acute infection, as previously hypothesized.[58,59] Although complement dysregulation and thrombosis are known to drive severe COVID-19,[54,60] this has not been extensively confirmed in long COVID and results from the HEAL-COVID study do not support the use of anticoagulation to enhance recovery after hospitalisation.[61]

Multiple mechanisms for long COVID have been suggested including autoimmunity, thrombosis, vascular dysfunction, SARS-CoV-2 persistence and latent virus reactivation.[6,54,62] The protein signatures we observed are consistent with these mechanisms, which can all result in myeloid inflammation and complement activation.[31,63] Given the heterogeneity of long COVID and the possibility of co-existing or multiple aetiologies, our work demonstrates a possible final common pathway between symptom groups that might be targetted by therapies and supports the rationale for several drugs which are currently under trial (Supplementary Table 4). Specifically, steroids, IL-1 antagonists (e.g. Anakinra), JAK inhibitors, cannabinoids and colchicine can suppress myeloid inflammation and/or vascular inflammation triggered by complement, and have acceptable safety profiles when used in other diseases.

Our findings of elevated anti-S and -N IgG in long COVID individuals strengthen the possibility of viral persistence, providing further rationale for current trials such as RECOVERY-VITAL using long courses of Paxlovid (Supplementary Table 4). However, it should be considered that combination antivirals may be required to successfully clear chronic SARS-CoV-2 infection.

This is the first study to identify protective mediators, finding an association between sCD58 and recovery. Given the role of sCD58 in supressing monocyte-lymphocyte interactions,[46,64] this finding adds strength to our conclusions that myeloid inflammation is central to the biology of long COVID. However, although the effect size of sCD58 estimated by PLR was substantial, this was variable when applying univariate and PLS analyses. Therefore, the protective role of sCD58 after COVID-19 requires confirmation in further studies.

We found differences in the proteome of adults with GI and cognitive symptoms. Our findings are supported by previous smaller studies demonstrating inconsistent associations with different symptoms. Two studies have suggested distinct mechanisms might underly neurological and cognitive symptoms,[6,35] while another identified autoantibodies specific to those with GI symptoms.[20] We did not measure autoantibodies but did identify markers of brain-gut axis disturbance associated with GI symptoms. The brain-gut axis is mediated by autonomic pathways including the vagus nerve, dysfunction of which has been implicated in autoimmune disease as well as long COVID.[2,65,66] Monocyte infiltration of the vagus nerve has been identified in acute COVID-19, and the myeloid signatures we observed might suggest this can also occur in long COVID.[67] We found specific signatures suggestive of neurodegeneration in those with cognitive impairment. This could support the previous suggestion of EBV reactivation in neurological long COVID,[6] since EBV infection has been shown to increase the risk of neurodegenerative disease[68] which could result from reactivation in the CNS. One recent study also found that patients with cognitive impairment

exhibited distinct and 'non-inflammatory' proteomic profiles consistent with our findings that proteins such as NFASC and SPON-1 were most strongly associated with cognitive symptoms.[35] Finally, our observations in individuals with anxiety/depression suggest that mechanisms for these symptoms are similar to those in non-COVID depression, also associated with IL-1 and myeloid inflammation.[69,70] Overall, our findings suggest that multiple and diverse processes associate with different long COVID symptoms and clinical trials may need to account for this when selecting patients.

Cardiorespiratory symptoms did not relate to upper respiratory tract inflammation. Due to the difficulties inherent in lower airway sampling, we measured nasal mediators, which can reflect lower airway inflammation in some scenarios.[71] However, this similarity between the upper and lower airway may not be true for long COVID, where isolated lung inflammation could contribute to symptoms.[12] Better understanding of the source of long COVID inflammation is important; systemic immunosuppression can have adverse side-effects and localised treatments such as inhaled corticosteroids might be preferable, if lung inflammation is confirmed.[72]

Not all immunological studies of long COVID support our findings. One proteomic study of 55 individuals did demonstrate IL-1 inflammation but found that TNF and anti-viral signatures dominated.[9] Notably, most individuals from that study experienced mild (WHO 2-3) disease and mechanisms may differ in long COVID after hospitalisation. Other studies have suggested vasculoproliferative processes and metabolic disturbance, but these studies used uninfected healthy controls for comparison and cannot distinguish between long COVID-specific phenomena and residual post-COVID inflammation.[5,11] Importantly, one study found no association between immune cell activation and long COVID.[10] However, this study only examined 63 adults with long COVID, 3 months after infection and did not measure IL-1 or monocyte function. Examining patients early after infection may not identify subtle differences since residual inflammation from the acute infection may dominate. A strength of our study is the examination of a large cohort experiencing different symptoms, several months after hospitalisation.

Our study has limitations. By design, we sought to identify inflammatory markers underlying long COVID symptom groups rather than identify biomarkers or prognostic signatures. This led to our primary use of PLR, the accuracy of which was appropriate for assessing associations between inflammatory markers and outcome (Supplementary). PLS and univariate analyses were used to support the reliability of PLR findings but were not used to draw conclusions. PLR enables relationships between mediators to be accounted for without false discovery.[73–75] In highly correlated data such as ours, PLR has been shown to consistently outperform PLS, which tends to overestimate coefficient variance.[34,76] Cognitive Impairment was the smallest group (n=65) resulting in a higher classification error for this analysis, however the final sensitivity of the model was 98%. Nonetheless, given the size of this group, larger studies would be useful to confirm our findings.

We used the validated WHO clinical progression score to classify the severity of acute infection.[77] Despite being a hospitalised cohort, the WHO progression scores indicated that individuals experienced a range of COVID-19 severities, including those who did not require oxygen (WHO class 3-4; Table 1). We did not find an association between acute COVID-19 severity and long COVID symptoms, suggesting that this did not influence the inflammatory profiles observed and our findings are consistent with those of a smaller study examining the proteome of 97 long COVID individuals, most of whom were not hospitalised during acute infection, equivalent to WHO 2-3 severity.[35] However, since we did not study non-hospitalised cases or cases of WHO 2 severity, our findings may only apply to a subset of long COVID cases. Many develop long COVID without requiring hospitalisation for acute infection, and given the heterogeneity of long COVID,[2] it is important that our findings are further validated in large cohorts of this type.

It is challenging to distinguish between long COVID and 'post-hospital syndrome' in our cohort. Whilst, the proteome of individuals with persistent symptoms after hospitalisation has not been well-characterised, the patterns of inflammation we observed were not similar to the typical 'immunosuppressive' profiles seen after severe sepsis or the antiviral profiles seen in post-ebola syndrome, suggesting our findings are specific to long COVID.[78–80]

The odds ratios we report are small, but the relationships identified were consistent across alternative methods of analysis and when using different long COVID definitions. The small effect sizes can be partly explained by the model used, which shrinks

correlated mediator coefficients towards each other to account for combined effects and prevent colinear inflation.[75] Thus, the effect sizes do not necessarily diminish the potential mechanistic significance of our findings since inflammatory proteins are expected to mediate effects in combination. The small effect sizes may also result from measurement of plasma mediators which likely underestimates the degree of inflammation in tissues.

The Olink platform has been extensively validated against other immunoassays;[81] we were able to compare measurements against alternative immunoassays for some mediators (Extended Data Fig. 8). In these instances, we saw good correlations between alternative methodologies but acknowledge that our analysis is largely reliant on the Olink platform.

Long COVID is poorly defined and presentations are heterogenous, making mechanistic studies challenging.[18] There is no gold standard tool for diagnosis and we aimed to use objective and validated measures of symptoms where feasible. A strength of this study is the mapping of inflammatory profiles onto common symptoms, which may support phenotyping. However, we were not able to encompass all long COVID symptoms (over 200 have been described)[2,18] and the use of MoCA scores to define cognitive impairment does not account for individuals with subjective 'brain fog'. We did not identify any patients who were unrecovered but were not assigned to a symptom group, indicating that no long COVID cases were excluded or misclassified.

Our findings of higher SARS-CoV-2-specific antibody responses in individuals with persistent symptoms support those of several studies.[13,20,35,52] Though we did not find direct evidence of SARS-CoV-2 in sputum, we were not able to exclude persistence of the type shown in studies of tissue from the GI tract or lung.[52,53] We were also not able to replicate observations of circulating S antigen in patients with long COVID.[82] Alternatively, dysregulation of the adaptive immune compartment, including autoimmunity might explain our findings in the absence of persistent virus, suggested by one recent report.[35] This is further supported by one study finding no difference in the early adaptive immune response to SARS-CoV-2 in those who later develop long COVID, making viral persistence through immune escape less plausible.[83] Notably, cohorts from both studies included mostly mild cases of acute COVID-19 (57/97[59%] WHO 2-3 and 76/136[56%] WHO 1-3 respectively). These varied reports highlight the diversity of mechanisms associated with long COVID, which may underpin the heterogeneity of long COVID symptoms. Our cohort of patients after hospital discharge may not fully represent this spectrum of disease.

In conclusion, we found markers of myeloid inflammation and complement activation associated with long COVID symptoms in a large cohort of individuals who were previously hospitalised for COVID-19 (Fig. 5). However, distinctive inflammatory patterns were seen in those with cognitive impairment and GI symptoms. Our findings support the use of immunomodulatory agents in therapeutic trials but demonstrate the need to consider the distinct inflammatory phenotypes between long COVID symptom groups.

Methods (Online)

Study design and Ethics

After hospitalisation for COVID-19, adults who had no co-morbidity resulting in a prognosis of less than 6 months were recruited to the PHOSP-COVID study. Patients that had been hospitalised between February 2020 and January 2021 were recruited. Both sexes were recruited and gender was self-reported. Written informed consent was obtained from all patients. Ethical approvals for the PHOSP-COVID study were given by Leeds West Research Ethics Committee (20/YH/0225).

Symptom data and samples were prospectively collected from individuals approximately 6 months after hospitalisation (Fig. 1A), via the PHOSP-COVID multicentre UK study.[16] Data relating to patient demographics and acute admission were collected via the International Severe Acute Respiratory and Emerging Infection Consortium World Health Organization Clinical Characterisation Protocol UK (ISARIC4C study; IRAS260007/IRAS126600).[84] Adults hospitalised during the SARS-CoV-2 pandemic were systematically recruited into ISARIC4C. Written informed consent was obtained from all patients. Ethical approval was given by the South Central–Oxford C Research Ethics Committee in England (reference: 13/SC/0149), Scotland A Research Ethics Committee (20/SS/0028) and World Health Organization Ethics Review Committee (RPC571 and RPC572l; 25 April 2013).

Data were collected to account for variables affecting symptom outcome, via hospital records and self-reporting. Acute disease severity was classified according to the WHO Clinical Progression score: Class 3-4 required no oxygen, Class 5 required oxygen therapy, Class 6 required non-invasive ventilation or high-flow nasal oxygen), Class 7-9 were managed in critical care.[77] Clinical data were used to place patients into 6 categories: 'Recovered', 'GI', 'Cardiorespiratory', 'Fatigue', 'Cognitive impairment' and 'Anxiety/depression' (Supplementary Table 5). Patient reported symptoms and validated clinical scores were used including: MRC breathlessness score, dyspnoea-12 score, FACIT score, PHQ-9 and GAD-7. Responses to symptom questionnaires about chest pain and palpitations were also used. Cognitive impairment was defined as a Montreal Cognitive Assessment (MoCA) score <26. GI symptoms were defined as answering 'Yes' to the presence of at least two of the listed symptoms. 'Recovered' was defined by self-reporting. Patients were placed in multiple groups if they experienced a combination of symptoms.

Matched nasal fluid and sputum samples were prospectively collected from a subgroup of convalescent patients approximately 6 months after hospitalisation via the PHOSP-COVID study. Nasal and bronchoalveolar lavage fluid (BALF) collected from healthy volunteers prior to the COVID-19 pandemic were used as controls (Supplementary Table 3). Written consent was obtained for all individuals and ethical approvals were given by London-Harrow Research Ethics Committee (13/LO/1899) for the collection of nasal samples and the Health Research Authority London–Fulham Research Ethics Committee (IRAS Project ID 154109; references 14/LO/1023, 10/H0711/94, and 11/LO/1826) for BALF samples.

Procedures

EDTA plasma was collected from whole blood taken by venepuncture and frozen at -80°C as previously described.[14,15] Nasal fluid was collected using a NasosorptionTM FX·I device (Hunt Developments UK Ltd), which uses a synthetic absorptive matrix to collect concentrated nasal fluid. Samples were eluted and stored as previously described.[85] Sputum samples were collected via passive expectoration and frozen at -80°C without addition of buffers. Sputum samples from convalescent individuals were compared to BALF from healthy SARS-CoV-2 naïve controls, collected before the pandemic. BALF samples were used to act as a comparison for lower respiratory tract samples since passively expectorated sputum from healthy SARS-CoV-2 naïve individuals was not available. BALF samples were obtained by instillation and recovery of up to 240 ml of normal saline via a fibreoptic bronchoscope. BALF was filtered through 100μM strainers into sterile 50ml Falcon tubes, then centrifuged for 10 minutes at 400g at 4°C. The resulting supernatant was transferred into sterile 50ml Falcon tubes and frozen at -80°C until use. The full methods for BALF collection and processing have been described.[86,87]

Immunoassays

To determine inflammatory signatures which associated with symptom outcomes, plasma samples were analysed on an Olink Explore 384 Inflammation panel (Uppsala, Sweden)[15]. Supplementary Table 6 (Appendix 1) lists all analytes measured. To ensure the validity of results, samples were run in a single batch with use of negative controls, plate controls in triplicate, and repeated measurement of patient samples between plates in duplicate. Samples were randomized between plates according to site and sample collection date. Randomization between plates was blind to long COVID/ Recovered outcome. Data were first normalized to an internal extension control that was included in each sample well. Plates were standardized by normalizing to inter-plate controls, run in triplicate on each plate. Each plate contained a minimum of 4 patient samples which were duplicates on another plate, these duplicate pairs allowed any plate to be linked to any other through the duplicates. Data were then intensity normalized across all cohort samples. Finally, Olink results underwent QC processing and samples or analytes that did not reach QC standards were excluded. Final normalized relative protein quantities were reported as log2 normalized protein expression (NPX) values.

To further validate our findings, we have compared our results to data previously generated from contemporaneously collected plasma from a subset of participants included in our study, using conventional electrochemiluminescence (ECL; MSD) or bead-based multiplex assays. Selecting analytes that were measured by both assays, we found good agreement between the methods (Extended Data Fig. 8). Notably, like most -omics platforms Olink measures relative quantities so perfect agreement with conventional assays that measure absolute concentrations is not expected.

Sputum samples were thawed prior to analysis and sputum plugs were extracted with the addition of 0.1% DTT creating a 1 in 2 sample dilution, as previously described.[88] SARS-CoV-2 Spike (S) and Nucleocapsid (N) proteins were measured by electrochemiluminescence S-plex assay at a fixed dilution of 1 in 2 (Mesoscale Diagnostics, Rockville, Maryland, USA), as per the manufacturers protocol.[89] Control BALF samples were thawed and measured on the same plate, neat. The S-plex assay is highly sensitive in detecting viral antigen in respiratory tract samples.[90]

Nasal cytokines were measured by ECL (MSD) and Luminex bead multiplex assays (Biotechne, Minneapolis, United States). The V-plex pro-inflammatory 1 kit (MSD), R-plex custom kit (MSD) and Human Premixed Multi-analyte custom kit (Biotechne) were used. Nasal samples were analysed at a fixed dilution of 1 in 2 using the R-plex and Luminex assays, and neat using the proinflammatory panel 1 kit. MSD plates were measured on a MESO QuickPlex SQ 120 Reader (MSD) and Luminex plates on a BioPlex200 instrument (Bio-Rad, UK). Cytokine concentrations were calculated using a reference standard and assigned pg/mL. All values at or below the lower limit of detection (LLOD) were replaced with LLOD. All values at or above the upper limit of detection (ULOD) were replaced with ULOD. The full methods and list of analytes are detailed in the Supplementary Materials.

Statistics

To determine protein signatures which associated with each symptom outcome, a ridge penalised logistic regression (PLR) was used. PLR shrinks coefficients to account for combined effects within high dimensional data, preventing false discovery whilst managing multicollinearity.[34] Thus, PLR was chosen *a priori* as the most appropriate model to assess associations between a large number of explanatory variables (which may work together to mediate effects) and symptom outcome.[34,74,75,91] In keeping with our aim to perform an unbiased exploration of inflammatory process, the model alpha was set to zero, facilitating regularisation without complete penalisation of any mediator. This enabled review of all possible mediators that might associate with long COVID.[75]

A 50 repeats 10-fold nested cross-validation was used to select the models with lowest classification error and optimal lambda (Extended Data Fig. 9). The cognitive impairment model had the lowest AUC and highest classification error due to the size of this group (n=65) relative to recovered (n=250). Thus, the model was weighted to account for this imbalance and preventing bias towards classification as the majority class. The weighted model had a sensitivity of 0.98 indicating its validity. The metrics of the model are discussed in the Supplementary Materials.

Age, sex, acute disease severity and pre-existing comorbidities were included as covariates in the PLR analysis (Supplementary Table 1,2). Covariates were selected *a priori* using features reported to influence the risk of long COVID and inflammatory responses.[2,16,48,92] Ethnicity was not included since it has been shown not to predict symptom outcome in this cohort.[16] Individuals with missing data were excluded from the regression analysis. Each symptom group was compared to the 'Recovered' group. The model coefficients of each covariate were converted into odds ratios for each outcome and visualised in a forest plot, after removing variables associated with regularised Odds Ratios (OR) between 0.98 and 1.02 and in cases where most variables fell outside of this range, using mediators associated with the highest decile of coefficients either side of this range. This enabled exclusion of mediators with effect sizes that were unlikely to have clinical or mechanistic significance since the ridge PLR shrinks and orders coefficients according to their relative importance rather than making estimates with standard error. Thus, confidence intervals cannot be appropriately derived from PLR and forest plot error bars were calculated using the median accuracy of the model generated by the nested cross-validation. To verify observations made through PLR analysis we also performed an unadjusted PLR, an unadjusted logistic regression and a Partial Least Squares (PLS) analysis. Univariate analyses using Wilcoxon signed rank test was also performed (Supplementary Table 6; Appendix 1). Analyses were performed in R version 4.2.0 using *'lme4', 'caret', 'glmnet', 'mdatools'* and *'ggplot2'* packages.

To further investigate the relationship between proteins elevated in each symptom group, we performed a correlation network analysis using Spearman's rank correlation coefficient and false discovery rate thresholding. The mediators visualised in the PLR forest plots, which were associated with Cardiorespiratory symptoms, Fatigue, Anxiety/Depression GI symptoms and Cognitive impairment were used respectively. Analyses were performed in R version 4.2.0 using '*bootnet'* and '*qgraph'* packages.

To determine if differences in protein levels between men and women related to hormonal differences, we divided each symptom group into pre-menopausal and post-menopausal groups using an age cut-off of 50 years old. Differences between sexes in each group were determined using the Wilcoxon-signed rank test. To understand if antigen persistence contributed to inflammation in adults with long COVID, the median viral antigen concentration from sputum/BALF samples and cytokine concentrations from nasal samples were compared using the Wilcoxon signed rank test. All tests were two-tailed and statistical significance was defined as a *p*-value<0.05 after adjustment for false discovery rate (q-value=0.05). Analyses were performed in R version 4.2.0 using *'ggpubr'* and *'ggplot2'* packages.

Declarations

Data sharing

The PHOSP-COVID protocol, consent form, definition and derivation of clinical characteristics and outcomes, training materials, regulatory documents, information about requests for data access, and other relevant study materials are available online: https://phosp.org/resource/. Access to these materials can be granted by contacting phosp@leicester.ac.uk and Phospcontracts@leicester.ac.uk.

The ISARIC4C protocol, data sharing and publication policy are available at https://isaric4c.net. ISARIC4C's Independent Data and Material Access Committee welcomes applications for access to data and materials (https://isaric4c.net).

All data and code used in this study is available within ODAP and accessible under reasonable request. Data access criteria and information about how to request access is available online: https://phosp.org/resource/. If criteria are met and a request is made, access can be gained by signing the eDRIS user agreement.

Acknowledgements

This research used data assets made available by Outbreak Data Analysis Platform (ODAP) as part of the Data and Connectivity National Core Study, led by Health Data Research UK in partnership with the Office for National Statistics and funded by UK Research and Innovation (grant ref MC_PC_20058).

This work is supported by the following grants: The PHOSP-COVD study is jointly funded by UK Research and Innovation and National Institute of Health and Care Research (grant references: MR/V027859/1 and COV0319). ISARIC4C is supported by grants from the National Institute for Health and Care Research (award CO-CIN-01) and the Medical Research Council (grant MC_PC_19059) Liverpool Experimental Cancer Medicine Centre provided infrastructure support for this research (grant reference: C18616/A25153). Other grants which have supported this work include: the UK Coronavirus Immunology Consortium [funder reference:1257927], the Imperial Biomedical Research Centre (NIHR Imperial BRC, grant IS-BRC-1215-20013), the Health Protection Research Unit (HPRU) in Respiratory Infections at Imperial College London and NIHR HPRU in Emerging and Zoonotic Infections at University of Liverpool, both in partnership with Public Health England, [NIHR award 200907], Wellcome Trust and Department for International Development [215091/Z/18/Z], Health Data Research UK (HDR UK) [grant code: 2021.0155], Medical Research Council [grant code: MC_UU_12014/12], and NIHR Clinical Research Network for providing infrastructure support for this research.

FL is supported by an MRC clinical training fellowship [award MR/W000970/1]. CE is funded by NIHR [grant P91258-4]. LPH is supported by Oxford NIHR Biomedical Research Centre. AART is supported by a BHF Intermediate Clinical Fellowship (FS/18/13/33281). SLRJ receives support from UKRI, GCRF, Rosetrees Trust, BHIVA, EDCTP, Globvac. JDC has grants from AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Gilead Sciences, Grifols, Novartis and Insmed. RAE holds a NIHR Clinician Scientist Fellowship (CS-2016-16-020). AH is currently supported by UK Research and Innovation. NIHR and NIHR Manchester BRC. BR receives support from BHF Oxford Centre of Research Excellence, NIHR Oxford BRC and MRC. DGW is supported by an NIHR Advanced Fellowship. AH has received support from MRC and for the Coronavirus Immunology Consortium (MR/V028448/1). LVW has received support from UKRI, GSK/Asthma + Lung UK and NIHR for this study. MGS has received

support from NIHR UK, MRC UK and Health Protection Research Unit in Emerging & Zoonotic Infections, University of Liverpool. JKB is supported by the Wellcome Trust (223164/Z/21/Z) and UKRI (MC_PC_20004, MC_PC_19025, MC_PC_1905, MRNO2995X/1, and MC_PC_20029). LT is supported by the Wellcome Trust [clinical career development fellowship grant number 205228/Z/16/Z], the Centre of Excellence in Infectious Diseases Research (CEIDR) and the Alder Hey Charity. LT is also supported by U.S. Food and Drug Administration Medical Countermeasures Initiative contract 75F40120C00085.PJMO is supported by a NIHR Senior Investigator Award [award 201385]. The funders were not involved in the study design, interpretation of data or writing of this manuscript. The views expressed are those of the authors and not necessarily those of the DHSC, DID, NIHR, MRC, the Wellcome Trust, UK-HAS, the National Health Service, or the Department of Health.

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 PHOSP-COVID study at all of the 65 NHS trusts/Health boards and 25 research institutions across the UK, as well as those who contributed to setting up and delivering the ISARIC4C study at 305 NHS trusts/ Health boards. We also thank all the supporting staff at the NIHR Clinical Research Network, Health Research Authority, Research Ethics Committee, Department of Health and Social Care, Public Health Scotland, and Public Health England. We thank Kate Holmes at the NIHR Office for Clinical Research Infrastructure (NOCRI) for her support in coordinating the charities group. The PHOSP-COVID industry framework was formed to provide advice and support in commercial discussions, and we thank the Association of the British Pharmaceutical Industry as well NOCRI for coordinating this. We are very grateful to all the charities that have provided insight to the study: Action Pulmonary Fibrosis, Alzheimer's Research UK, Asthma + Lung UK, British Heart Foundation, Diabetes UK, Cystic Fibrosis Trust, Kidney Research UK, MQ Mental Health, Muscular Dystrophy UK, Stroke Association Blood Cancer UK, McPin Foundations, and Versus Arthritis. We thank the NIHR Leicester Biomedical Research Centre patient and public involvement group and Long Covid Support. We would also like to thank Professor Golam Khandaker and Professor Dawn C. Newcomb who provided valuable feedback on this work. Figure 5 was created using Biorender.

Contributors

FL recruited participants, acquired clinical samples, analysed and interpreted data and co-wrote the manuscript, including all drafting and revisions. **CE** analysed and interpreted data and co-wrote this manuscript, including all drafting and revisions. **SF** and **MR** supported analysis and interpretation of data as well as drafting and revisions. **DS, JKS, SCM, SA, NM, JN, CK, OCL, OE, HJCM, ASh, ASi, MS, VCH, MT, NJG, NIL, CC** contributed to acquisition of data underlying this study. **LHW**, **AART, SLRJ, LH, OMK, DGW, TIdS** and **AH** made substantial contributions to conception/design and implementation of this work and/or acquisition of clinical samples for this work. They have supported drafting and revisions of the manuscript. **EH, JKQ** and **ABD** made substantial contributions to study design as well as data access, linkage and analysis. They have supported drafting and revisions of this work. **JDC, LPH, AH, BR, KP, MM, WG** made substantial contributions to conception and design of this work and have supported drafting and revisions of this work. **JKB** obtained funding for ISARIC4C, is ISARIC4C consortium co-lead, has made substantial contributions to conception and design of this work and has supported drafting and revisions of this work. **MGS** obtained funding for ISARIC4C, is ISARIC4C consortium co-lead, sponsor/protocol chief investigator, has made substantial contributions to conception and design of this work and has supported drafting and revisions of this work. **RAE and LVW** are co-leads of PHOSP-COVID, made substantial contributions to conception and design of this work, acquisition and analysis of data, and have supported drafting and revisions of this work. **CB** is the chief investigator of PHOSP-COVID and has made substantial contributions to conception and design of this work. **RST** and **LT** made substantial contributions to acquisition, analysis and interpretation of the data underlying this study and have contributed to drafting and revisions of this work. **PJMO** obtained funding for ISARIC4C, is ISARIC4C consortium co-lead, sponsor/protocol chief investigator, and has made substantial contributions to conception and design of this work. **RST** and **PJMO** have also made key contributions to interpretation of data and have co-written this manuscript. All authors have read and approve the final version to be published. All authors agree to accountability for all aspects of this work.

All investigators within ISARIC4C and the PHOSP-COVID consortia have made substantial contributions to the conception or design of this study and/or acquisition of data for this study. The full list of authors within these groups is available in the supplementary materials.

Declaration of interests

FL, CE, DS, JKS, SCM, CD, CK, NM, LN, EH, ABD, JKQ, LPH, KP, LH, OMK, SF, TIdS, DGW, RST and JKB have no conflicts of interest. AART receives speaker fees and support to attend meetings from Janssen Pharmaceuticals. SLRJ is on the data safety monitoring board for Bexero trial in HIV+ adults in Kenya. JDC is the deputy chief editor of ERS and receives consulting fees from AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Insmed, Janssen, Novartis, Pfizer and Zambon. AH is Deputy chair of NIHR Translational Research Collaboration (unpaid role). BR receives honoraria from Axcella therapeutics. RAE is co-lead of PHOSP-COVID and receives fees from Astrazenaca / Evidera for consultancy on Long Covid and from Astrazenaca for consultancy on digital health. RAE has received speaker fees from Boehringer in June 2021 and has held a role as European Respiratory Society Assembly 01.02 Pulmonary Rehabilitation secretary. RAE is on the American Thoracic Society Pulmonary Rehabilitation Assembly programme committee. LVW also receives funding from Orion pharma and GSK and holds contracts with Genentech and AstraZenaca. LVW has received consulting fees from Galapagos and Boehringer, is on the data advisory board for Galapagos and is Associate Editor for European Respiratory Journal. AH is a member of NIHR Urgent Public Health Group (June 2020-March 2021). MM is an applicant on the PHOSP study funded by NIHR/DHSC. MGS acts as an independent external and non-remunerated member of Pfizer's External Data Monitoring Committee for their mRNA vaccine program(s), is Chair of Infectious Disease Scientific Advisory Board of Integrum Scientific LLC, Greensboro, NC, USA and is director of MedEx Solutions Ltd and majority owner of MedEx Solutions Ltd and minority owner of Integrum Scientific LLC, Greensboro, NC, USA. MGS's institution has been in receipt of gifts from Chiesi Farmaceutici S.p.A. of Clinical Trial Investigational Medicinal Product without encumbrance and distribution of same to trial sites. MGS is a non-renumerated member of HMG UK New Emerging Respiratory Virus Threats Advisory Group (NERVTAG) and has previously been a non-renumerated member of SAGE. CB has received consulting fees and/or grants from GSK, AZ, Genentech, Roche, Novartis, Sanofi, Regeneron, Chiesi, Mologic and 4DPharma. LT has received consulting fees from MHRA and speak fees from Eisai Ltd. LT has a patent pending with ZikaVac. PJMO reports grants from the EU Innovative Medicines Initiative (IMI) 2 Joint Undertaking during the submitted work; grants from UK Medical Research Council, GlaxoSmithKline, Wellcome Trust, EU-IMI, UK, National Institute for Health Research, and UK Research and Innovation-Department for Business, Energy and Industrial Strategy; and personal fees from Pfizer, Janssen, and Seqirus, outside the submitted work.

References

1. Wise, J. Covid-19: WHO urges action as 17 million long covid cases are estimated in Europe. *BMJ* o2232 (2022) doi:10.1136/bmj.o2232.
2. Davis, H. E., McCorkell, L., Vogel, J. M. & Topol, E. J. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol* **21**, 133–146 (2023).
3. Antonelli, M., Pujol, J. C., Spector, T. D., Ourselin, S. & Steves, C. J. Risk of long COVID associated with delta versus omicron variants of SARS-CoV-2. *The Lancet* **399**, 2263–2264 (2022).
4. Mizrahi, B. *et al.* Long covid outcomes at one year after mild SARS-CoV-2 infection: nationwide cohort study. *BMJ* e072529 (2023) doi:10.1136/bmj-2022-072529.
5. Captur, G. *et al.* Plasma proteomic signature predicts who will get persistent symptoms following SARS-CoV-2 infection. *EBioMedicine* **85**, 104293 (2022).
6. Peluso, M. J. *et al.* Chronic viral coinfections differentially affect the likelihood of developing long COVID. *Journal of Clinical Investigation* **133**, (2023).
7. Phetsouphanh, C. *et al.* Immunological dysfunction persists for 8 months following initial mild-to-moderate SARS-CoV-2 infection. *Nat Immunol* **23**, 210–216 (2022).
8. Schultheiß, C. *et al.* The IL-1β, IL-6, and TNF cytokine triad is associated with post-acute sequelae of COVID-19. *Cell Rep Med* **3**, 100663 (2022).
9. Talla, A. *et al.* Persistent serum protein signatures define an inflammatory subcategory of long COVID. *Nat Commun* **14**, 3417 (2023).

10. Santopaolo, M. *et al.* Prolonged T-cell activation and long COVID symptoms independently associate with severe COVID-19 at 3 months. *Elife* **12**, (2023).
11. Iosef, C. *et al.* Plasma proteome of Long-COVID patients indicates HIF-mediated vasculo-proliferative disease with impact on brain and heart function. *J Transl Med* **21**, 377 (2023).
12. Scott, N. A. *et al.* Monocyte migration profiles define disease severity in acute COVID-19 and unique features of long COVID. *European Respiratory Journal* 2202226 (2023) doi:10.1183/13993003.02226-2022.
13. Klein, J. *et al.* Distinguishing features of Long COVID identified through immune profiling. *Nature* (2023) doi:10.1038/s41586-023-06651-y.
14. Elneima, O. *et al.* Cohort Profile: Post-hospitalisation COVID-19 study (PHOSP-COVID). *medRxiv* 2023.05.08.23289442 (2023) doi:10.1101/2023.05.08.23289442.
15. Evans, R. A. *et al.* Clinical characteristics with inflammation profiling of long COVID and association with 1-year recovery following hospitalisation in the UK: a prospective observational study. *Lancet Respir Med* **10**, 761–775 (2022).
16. Evans, R. A. *et al.* Physical, cognitive, and mental health impacts of COVID-19 after hospitalisation (PHOSP-COVID): a UK multicentre, prospective cohort study. *Lancet Respir Med* **9**, 1275–1287 (2021).
17. Zhang, H. *et al.* Data-driven identification of post-acute SARS-CoV-2 infection subphenotypes. *Nat Med* **29**, 226–235 (2023).
18. Davis, H. E. *et al.* Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine* **38**, 101019 (2021).
19. Klein, J. *et al.* Distinguishing features of Long COVID identified through immune profiling. *medRxiv* 2022.08.09.22278592 (2022) doi:10.1101/2022.08.09.22278592.
20. Su, Y. *et al.* Multiple early factors anticipate post-acute COVID-19 sequelae. *Cell* **185**, 881-895.e20 (2022).
21. Muri, J. *et al.* Autoantibodies against chemokines post-SARS-CoV-2 infection correlate with disease course. *Nat Immunol* **24**, 604–611 (2023).
22. Houchen-Wolloff, L. *et al.* Joint patient and clinician priority setting to identify 10 key research questions regarding the long-term sequelae of COVID-19. *Thorax* **77**, 717–720 (2022).
23. National Institute for Health and Care Excellence (NICE), Scottish Intercollegiate Guidelines Network (SIGN) & Royal College of General Practitioners (RCGP). *COVID-19 rapid guideline: managing the long-term effects of COVID-19*. https://www.nice.org.uk/guidance/ng188/resources/covid19-rapid-guideline-managing-the-longterm-effects-of-covid19-pdf-51035515742 (2022).
24. Centers for Disease Control and Prevention. Long COVID or Post-COVID Conditions. *https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects/index.html#:~:text=Long%20COVID%20is%20broadly%20defined,after%20acute%20COVID%2D19%20infection.* (2023).
25. Peters, V. A., Joesting, J. J. & Freund, G. G. IL-1 receptor 2 (IL-1R2) and its role in immune regulation. *Brain Behav Immun* **32**, 1–8 (2013).
26. The, E. *et al.* Interleukin 38 alleviates aortic valve calcification by inhibition of NLRP3. *Proceedings of the National Academy of Sciences* **119**, (2022).
27. Luo, Z. *et al.* Monocytes augment inflammatory responses in human aortic valve interstitial cells via β2-integrin/ICAM-1-mediated signaling. *Inflammation Research* **71**, 681–694 (2022).
28. Bendall, L. J. & Bradstock, K. F. G-CSF: From granulopoietic stimulant to bone marrow stem cell mobilizing agent. *Cytokine Growth Factor Rev* **25**, 355–367 (2014).
29. Coombs, P. J., Graham, S. A., Drickamer, K. & Taylor, M. E. Selective Binding of the Scavenger Receptor C-type Lectin to Lewisx Trisaccharide and Related Glycan Ligands. *Journal of Biological Chemistry* **280**, 22993–22999 (2005).
30. Ma, Y. J. *et al.* Soluble Collectin-12 (CL-12) Is a Pattern Recognition Molecule Initiating Complement Activation via the Alternative Pathway. *The Journal of Immunology* **195**, 3365–3373 (2015).

31. Laursen, N. S. *et al.* Functional and Structural Characterization of a Potent C1q Inhibitor Targeting the Classical Pathway of the Complement System. *Front Immunol* **11**, (2020).
32. Dejanovic, B. *et al.* Complement C1q-dependent excitatory and inhibitory synapse elimination by astrocytes and microglia in Alzheimer's disease mouse models. *Nat Aging* **2**, 837–850 (2022).
33. The World Health Organization. Post COVID-19 condition (Long COVID). *https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition#:~:text=Definition,months%20with%20no%20other%20explanation.* (2022).
34. Firinguetti, L., Kibria, G. & Araya, R. Study of partial least squares and ridge regression methods. *Commun Stat Simul Comput* **46**, 6631–6644 (2017).
35. Woodruff, M. C. *et al.* Chronic inflammation, neutrophil activity, and autoreactivity splits long COVID. *Nat Commun* **14**, 4201 (2023).
36. Taquet, M. *et al.* Acute blood biomarker profiles predict cognitive deficits 6 and 12 months after COVID-19 hospitalization. *Nat Med* (2023) doi:10.1038/s41591-023-02525-y.
37. Luo, S., Hu, D., Wang, M., Zipfel, P. F. & Hu, Y. Complement in Hemolysis- and Thrombosis- Related Diseases. *Front Immunol* **11**, (2020).
38. Xue, G., Hua, L., Zhou, N. & Li, J. Characteristics of immune cell infiltration and associated diagnostic biomarkers in ulcerative colitis: results from bioinformatics analysis. *Bioengineered* **12**, 252–265 (2021).
39. He, T. *et al.* Integrative computational approach identifies immune-relevant biomarkers in ulcerative colitis. *FEBS Open Bio* **12**, 500–515 (2022).
40. Sundin, J. *et al.* Fecal chromogranins and secretogranins are linked to the fecal and mucosal intestinal bacterial composition of IBS patients and healthy subjects. *Sci Rep* **8**, 16821 (2018).
41. Kriebel, M., Wuchter, J., Trinks, S. & Volkmer, H. Neurofascin: A switch between neuronal plasticity and stability. *Int J Biochem Cell Biol* **44**, 694–697 (2012).
42. Woo, W.-M. *et al.* The C. elegans F-spondin family protein SPON-1 maintains cell adhesion in neural and non-neural tissues. *Development* **135**, 2747–2756 (2008).
43. Demydchuk, M. *et al.* Insights into Hunter syndrome from the structure of iduronate-2-sulfatase. *Nat Commun* **8**, 15786 (2017).
44. Parish, C. R. The role of heparan sulphate in inflammation. *Nat Rev Immunol* **6**, 633–643 (2006).
45. Zhang, Y., Liu, Q., Yang, S. & Liao, Q. CD58 Immunobiology at a Glance. *Front Immunol* **12**, (2021).
46. Hoffmann, J. C. *et al.* A soluble form of the adhesion receptor CD58 (LFA-3) is present in human body fluids. *Eur J Immunol* **23**, 3003–3010 (1993).
47. Bonilla, H. *et al.* Therapeutic trials for long COVID-19: A call to action from the interventions taskforce of the RECOVER initiative. *Front Immunol* **14**, (2023).
48. Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nat Rev Immunol* **16**, 626–638 (2016).
49. Real, F. G. *et al.* Lung function, respiratory symptoms, and the menopausal transition. *Journal of Allergy and Clinical Immunology* **121**, 72-80.e3 (2008).
50. Cheng, M. I. *et al.* The X-linked epigenetic regulator UTX controls NK cell-intrinsic sex differences. *Nat Immunol* **24**, 780–791 (2023).
51. Migeon, B. R. X-linked diseases: susceptible females. *Genetics in Medicine* **22**, 1156–1174 (2020).
52. Gaebler, C. *et al.* Evolution of antibody immunity to SARS-CoV-2. *Nature* **591**, 639–644 (2021).
53. Bussani, R. *et al.* Persistent SARS-CoV-2 infection in patients seemingly recovered from COVID-19. *J Pathol* **259**, 254–263 (2023).
54. Potere, N. *et al.* NLRP3 inflammasome and interleukin-1 contributions to COVID-19-associated coagulopathy and immunothrombosis. *Cardiovasc Res* (2023) doi:10.1093/cvr/cvad084.
55. Lammi, V. *et al.* Genome-wide Association Study of Long COVID. *medRxiv* (2023) doi:10.1101/2023.06.29.23292056.

56. Ismailova, A. *et al.* Identification of a forkhead box protein transcriptional network induced in human neutrophils in response to inflammatory stimuli. *Front Immunol* **14**, (2023).
57. Morgan, B. P. & Harris, C. L. Complement, a target for therapy in inflammatory and degenerative diseases. *Nat Rev Drug Discov* **14**, 857–877 (2015).
58. Nalbandian, A. *et al.* Post-acute COVID-19 syndrome. *Nat Med* **27**, 601–615 (2021).
59. Proal, A. D. & VanElzakker, M. B. Long COVID or Post-acute Sequelae of COVID-19 (PASC): An Overview of Biological Factors That May Contribute to Persistent Symptoms. *Front Microbiol* **12**, (2021).
60. Siggins, M. K. *et al.* Alternative pathway dysregulation in tissues drives sustained complement activation and predicts outcome across the disease course in COVID-19. *Immunology* **168**, 473–492 (2023).
61. Toshner, M. R. *et al.* Apixaban following discharge in hospitalised adults with COVID-19: Preliminary results from a multicentre, open-label, randomised controlled platform clinical trial. *medRxiv* (2022) doi:10.1101/2022.12.07.22283175.
62. Davis, H. E., McCorkell, L., Vogel, J. M. & Topol, E. J. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol* (2023) doi:10.1038/s41579-022-00846-2.
63. Beurskens, F. J., van Schaarenburg, R. A. & Trouw, L. A. C1q, antibodies and anti-C1q autoantibodies. *Mol Immunol* **68**, 6–13 (2015).
64. Zhang, Y., Liu, Q., Yang, S. & Liao, Q. CD58 Immunobiology at a Glance. *Front Immunol* **12**, 705260 (2021).
65. Lladós, G. *et al.* Vagus Nerve Dysfunction in the Post-COVID-19 Condition. *SSRN: https://ssrn.com/abstract=4479598* .
66. Bellocchi, C. *et al.* The Interplay between Autonomic Nervous System and Inflammation across Systemic Autoimmune Diseases. *Int J Mol Sci* **23**, 2449 (2022).
67. Woo, M. S. *et al.* Vagus nerve inflammation contributes to dysautonomia in COVID-19. *Acta Neuropathol* (2023) doi:10.1007/s00401-023-02612-x.
68. Levine, K. S. *et al.* Virus exposure and neurodegenerative disease risk across national biobanks. *Neuron* **111**, 1086-1093.e2 (2023).
69. Foley, É. M., Parkinson, J. T., Mitchell, R. E., Turner, L. & Khandaker, G. M. Peripheral blood cellular immunophenotype in depression: a systematic review and meta-analysis. *Mol Psychiatry* **28**, 1004–1019 (2023).
70. Dantzer, R., O'Connor, J. C., Freund, G. G., Johnson, R. W. & Kelley, K. W. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat Rev Neurosci* **9**, 46–56 (2008).
71. Jha, A. *et al.* Increased nasal mucosal interferon and CCL13 response to a TLR7/8 agonist in asthma and allergic rhinitis. *Journal of Allergy and Clinical Immunology* **147**, 694-703.e12 (2021).
72. Ramakrishnan, S. *et al.* Inhaled budesonide in the treatment of early COVID-19 (STOIC): a phase 2, open-label, randomised controlled trial. *Lancet Respir Med* **9**, 763–772 (2021).
73. Algamal, Z. Y. & Lee, M. H. Applying Penalized Binary Logistic Regression with Correlation Based Elastic Net for Variables Selection. *Journal of Modern Applied Statistical Methods* **14**, 168–179 (2015).
74. Greenwood, C. J. *et al.* A comparison of penalised regression methods for informing the selection of predictive markers. *PLoS One* **15**, e0242730 (2020).
75. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* **33**, 1–22 (2010).
76. Xia, Y. Correlation and association analyses in microbiome study integrating multiomics in health and disease. in 309–491 (2020). doi:10.1016/bs.pmbts.2020.04.003.
77. Marshall, J. C. *et al.* A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis* **20**, e192–e197 (2020).
78. Wiedemann, A. *et al.* Long-lasting severe immune dysfunction in Ebola virus disease survivors. *Nat Commun* **11**, 3730 (2020).
79. Hiser, S. L., Fatima, A., Ali, M. & Needham, D. M. Post-intensive care syndrome (PICS): recent updates. *J Intensive Care* **11**, 23 (2023).

80. Voiriot, G. *et al.* Chronic critical illness and post-intensive care syndrome: from pathophysiology to clinical challenges. *Ann Intensive Care* **12**, 58 (2022).
81. Wik, L. *et al.* Proximity Extension Assay in Combination with Next-Generation Sequencing for High-throughput Proteome-wide Analysis. *Molecular & Cellular Proteomics* **20**, 100168 (2021).
82. Swank, Z. *et al.* Persistent Circulating Severe Acute Respiratory Syndrome Coronavirus 2 Spike Is Associated With Post-acute Coronavirus Disease 2019 Sequelae. *Clinical Infectious Diseases* **76**, e487–e490 (2023).
83. Altmann, D. M. *et al.* Persistent symptoms after COVID-19 are not associated with differential SARS-CoV-2 antibody or T cell immunity. *Nat Commun* **14**, 5139 (2023).
84. Docherty, A. B. *et al.* Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* m1985 (2020) doi:10.1136/bmj.m1985.
85. Liew, F. *et al.* SARS-CoV-2-specific nasal IgA wanes 9 months after hospitalisation with COVID-19 and is not induced by subsequent vaccination. *EBioMedicine* **87**, 104402 (2023).
86. Ascough, S. *et al.* Divergent age-related humoral correlates of protection against respiratory syncytial virus infection in older and young adults: a pilot, controlled, human infection challenge model. *Lancet Healthy Longev* **3**, e405–e416 (2022).
87. Guvenel, A. *et al.* Epitope-specific airway-resident CD4+ T cell dynamics during experimental human RSV infection. *Journal of Clinical Investigation* **130**, 523–538 (2019).
88. Higham, A. *et al.* Leukotriene B4 levels in sputum from asthma patients. *ERJ Open Res* **2**, 00088–02015 (2016).
89. MSD. SARS-CoV-2 Spike Kit (product insert). *https://www.mesoscale.com/~/media/files/product%20inserts/s-plex%20sars-cov-2%20spike%20kit%20product%20insert.pdf* (2023).
90. Ren, A. *et al.* Ultrasensitive assay for saliva-based SARS-CoV-2 antigen detection. *Clinical Chemistry and Laboratory Medicine (CCLM)* **60**, 771–777 (2022).
91. Breheny, P. & Huang, J. Penalized methods for bi-level variable selection. *Stat Interface* **2**, 369–380 (2009).
92. Thwaites, R. S. *et al.* Inflammatory profiles across the spectrum of disease reveal a distinct role for GM-CSF in severe COVID-19. *Sci Immunol* **6**, (2021).

Table

Table 1. Cohort demographics. The demographics of each symptom group and recovered controls are shown. The WHO clinical progression scale was used to classify acute COVID-19 severity: Class 3-4=no oxygen requirement, Class 5=oxygen therapy, Class 6=non-invasive ventilation or high-flow nasal oxygen and Class 7-9=Organ support. Differences between groups were compared using chi-squared, Kruskal-Wallis or ANOVA as appropriate. Data are n (%) or mean (SD). C-Reactive Protein (CRP) levels represent those measured contemporaneously with clinical data collected in this study.

		GI	Fatigue	Cardio-respiratory	Anxiety/ depression	Cognitive impairment	Recovered	*p*
Age at admission	Years (SD)	57.72 (11.48)	56.57 (11.07)	57.08 (11.37)	56.36 (10.84)	59.24 (12.82)	58.92 (13.72)	*p*=0.046 *
Sex	Female N (%)	68 (53%)	143 (47%)	161 (43%)	89 (45%)	24 (42%)	55 (27%)	*p*<0.001 ****
Ethnicity	White	110	300	331	193	50	197	*p*=0.09 NS
	South Asian	14	26	38	16	7	46	
	Black	8	16	25	11	7	10	
	Mixed/Other	8	24	22	16	7	17	
WHO Clinical Progression Scale	Class 3-4	41	83	88	45	18	45	*p*=0.28 NS
	Class 5	45	107	124	74	21	115	
	Class 6	27	78	89	55	11	57	
	Class 7-9	27	98	115	62	21	50	
CRP	Mean (SD)	5.33 (5.42)	5.47 (7.17)	5.17 (6.82)	5.79 (8.12)	4.58 (5.78)	4.75 (10.38)	*p*=0.76 NS
Length of hospitalization	Days (SD)	12.04 (14.3)	14.59 (18.41)	15.39 (19.96)	14.57 (17.76)	14.95 (16.01)	12.5 (15.73)	*p*=0.0047 **
Steroid†	% Yes	34%	35%	37%	38%	33%	29%	*p*=0.294 NS
Remdesivir†	% Yes	4%	3%	4%	2%	3%	3%	*p*=0.725 NS
Comorbidities	Mean (SD)	2.9 (2.62)	2.675 (2.3)	2.553 (2.24)	2.911 (2.47)	2.493 (2.17)	1.554 (1.67)	*p*= <0.0001 ****

†Denotes treatment given during acute illness.

Figures

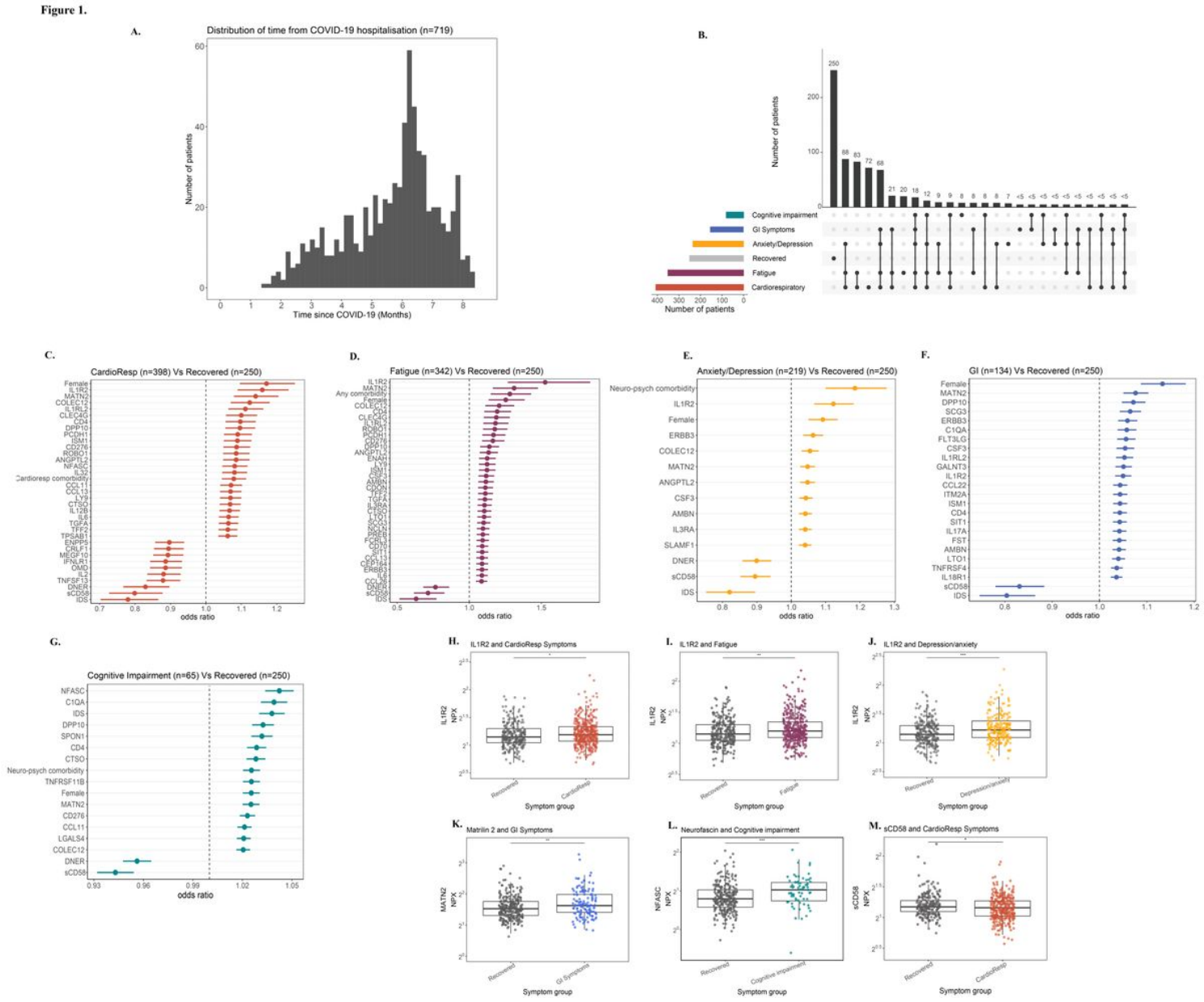

Figure 1

Long COVID symptoms and associated protein signature. Clinical data and samples were collected from 719 patients; the average time between hospital discharge and sample collection was 6 months. The full distribution of dates is shown for all patients used in the primary analysis (**A**), using the CDC and NICE criteria for long COVID (>4 weeks after COVID-19). Patients were grouped according symptoms reported (**B**). The horizontal, coloured bars represent the number of patients in each symptom group. The vertical black bars represent the number of patients in each symptom combination group. To prevent patient identification, where less than 5 patients belong to a combination group, this has been represented as '<5'. Within the cohort 250 patients reported feeling recovered and were used as controls ("Recovered", grey horizontal bar). Forest plots of protein signatures associated with Cardiorespiratory symptoms (**C**), Fatigue (**D**), Anxiety/depression (**E**), GI symptoms (**F**) and Cognitive impairment (**G**). Odds ratios were derived from PLR coefficients adjusted for sex, age, acute disease severity and comorbidities. Confidence intervals cannot be derived from PLR and error bars represent the median accuracy of the model. The distribution of inflammatory proteins associated with the highest odds of each outcome and compared to protein levels in the recovered group are shown (**H–L**). The association between sCD58 and recovery compared to cardiorespiratory symptoms is shown (**M**). Median values were compared between groups using the Wilcoxon signed-rank test. * = p<0·05, ** = p<0·01, *** = p<0·001, ****= p<0·0001.

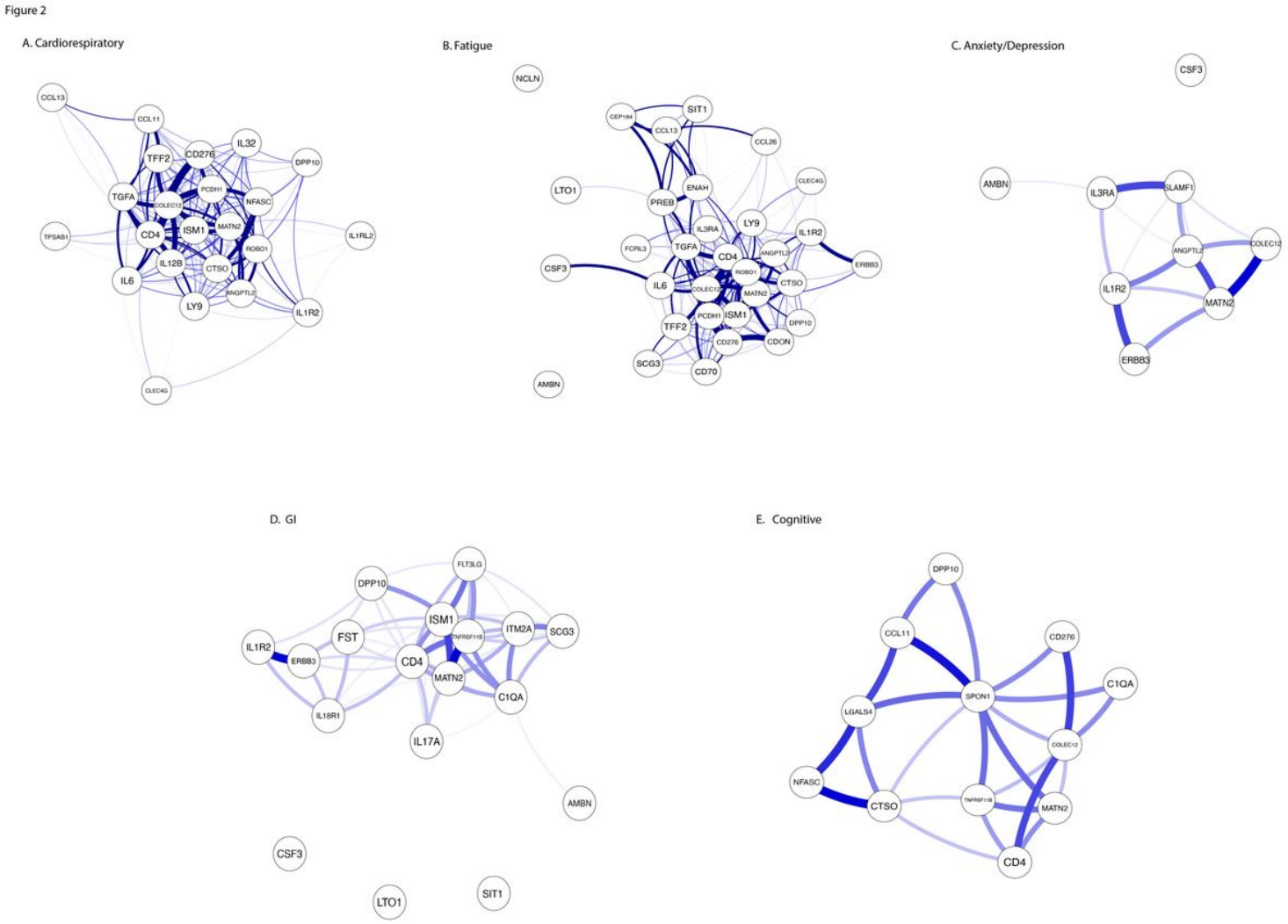

Figure 2

Relationship between mediators in each symptom group. A network analysis was performed using mediators associated with Cardiorespiratory symptoms (**A**), Fatigue (**B**), Anxiety/Depression (**C**), GI symptoms (**D**) and Cognitive Impairment (**E**). Each node corresponds to a protein mediator identified by the PLR. Edges (blue lines) were weighted according to the size of Spearman's rank correlation coefficient between proteins. All edges represent positive and significant correlations ($p<0.05$) after FDR adjustment.

Figure 3.

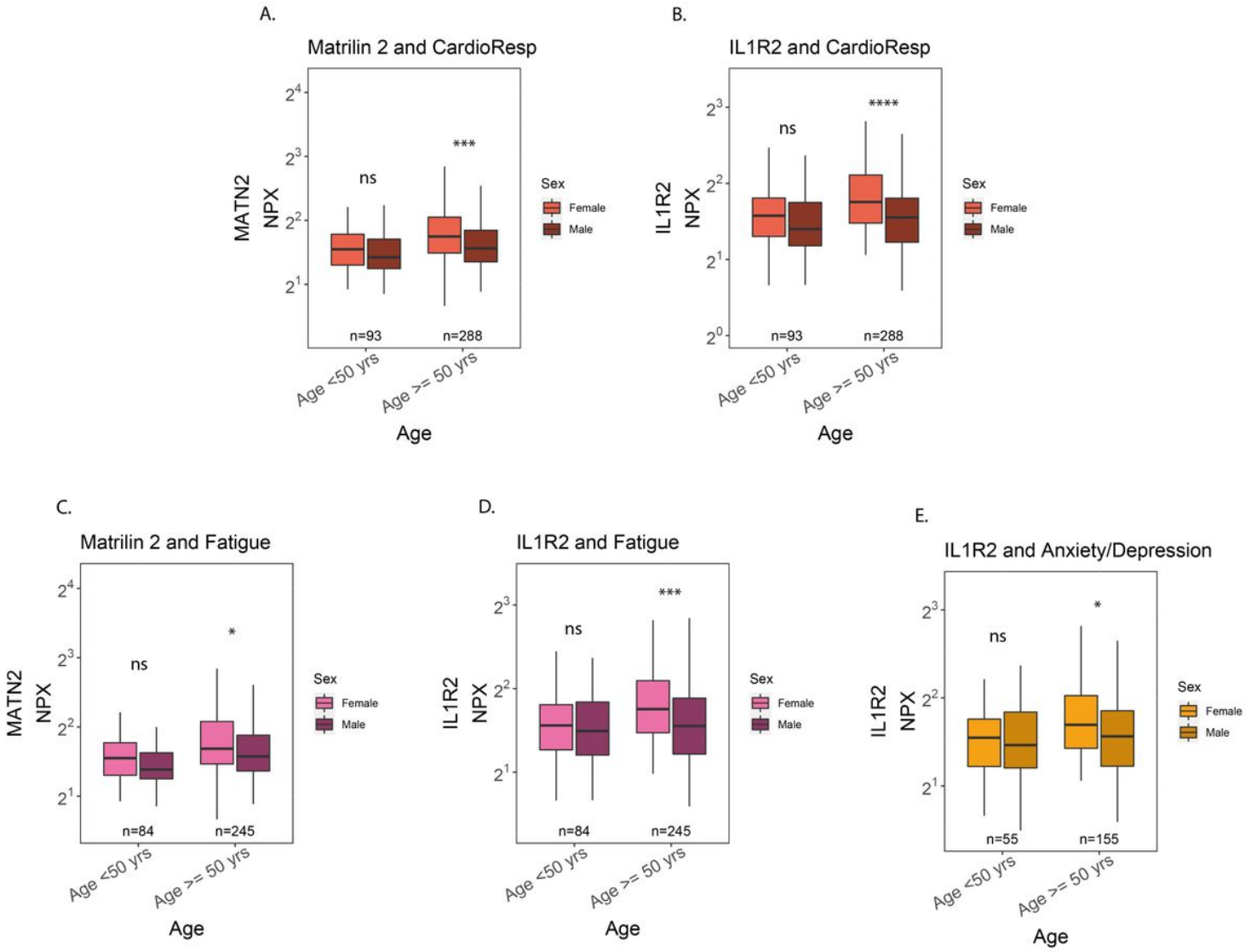

Figure 3

Inflammatory responses in men and women with long COVID. Patients were divided by age group and sex to understand differences between men and women with long COVID. Individuals were grouped according to the approximate age of menopause in women (50 years) to determine if oestrogen influences inflammation in long COVID. Levels of MATN2 (**A**) and IL-1R2 (**B**) in individuals with Cardiorespiratory symptoms are shown. Levels of MATN2 (**C**) and IL-1R2 (**D**) in individuals with Fatigue are shown. Levels of IL-1R2 within the Anxiety/Depression group are shown (**E**). Median values were compared between men and women using the Wilcoxon signed-rank test. * = p<0·05, ** = p<0·01, *** = p<0·001, ****= p<0·0001.

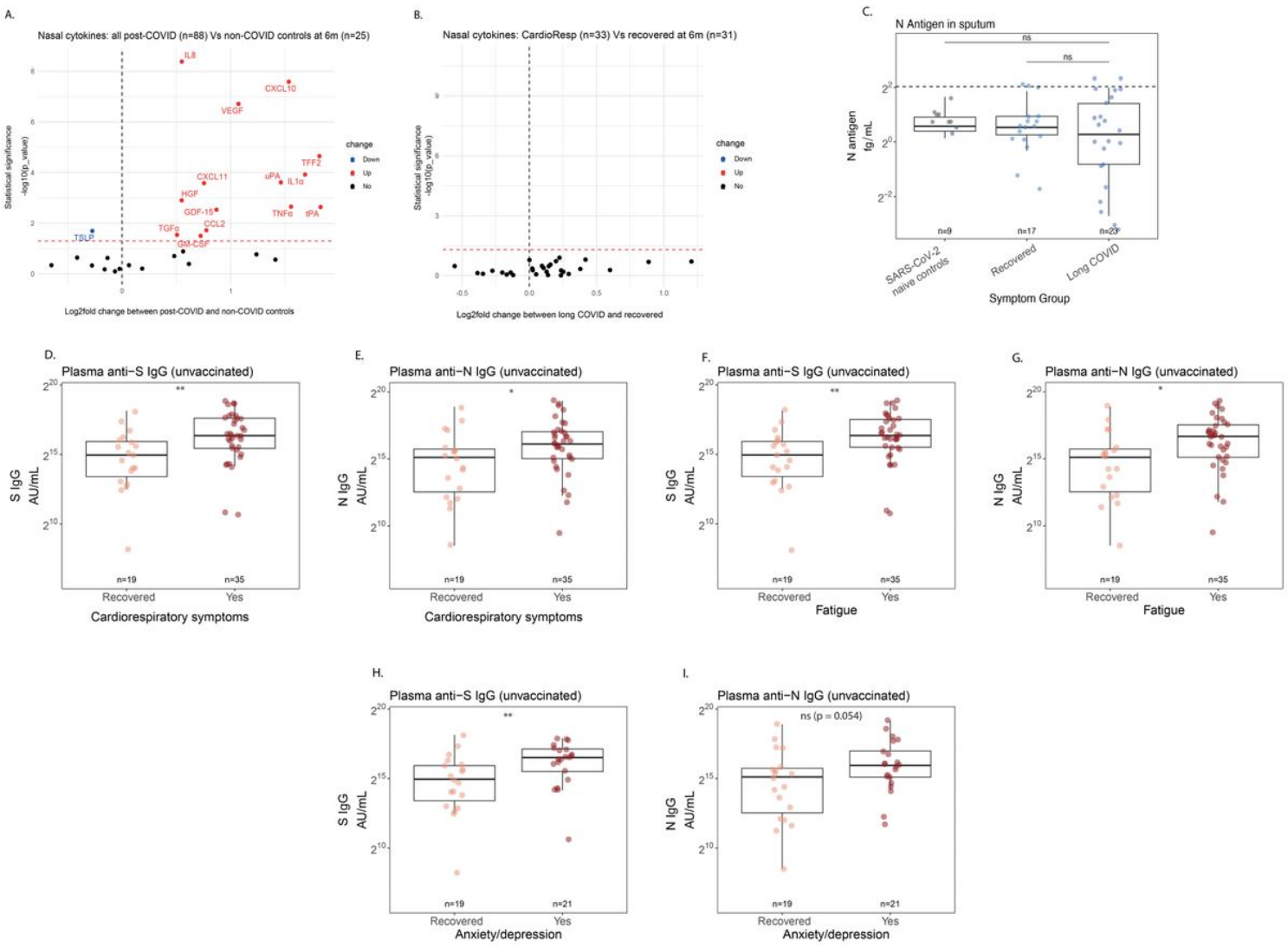

Figure 4

Sources of long COVID inflammation. Volcano plots showing the log2-fold change in nasal cytokine concentration between 88 post-COVID patients and 25 pre-pandemic healthy control samples (**A**) and between long COVID patients (n=33) experiencing Cardiorespiratory symptoms, and post-COVID patients who felt recovered (n=31) (**B**). The red values indicate increased cytokine levels between groups and a significant change after FDR adjustment ($p<0.05$). The blue values denote depressed cytokine levels and a significant change after FDR adjustment ($p<0.05$). Sputum was analysed for evidence of SARS-CoV-2 Nucleocapsid antigen (**C**) and compared to pre-pandemic BALF samples. The horizontal dashed line indicates the LLOD of the assay. Plasma Anti-S (**D**) and anti-N (**E**) IgG responses were measured in those with cardiorespiratory symptoms compared to recovered controls. Anti-S (**F**) and anti-N IgG (**G**) responses were compared in those with Fatigue and those recovered. Anti-S (**H**) and anti-N IgG (**I**) responses measured in those with Anxiety/Depression, compared to recovered controls. Median values were compared using the Wilcoxon signed-rank test. * = $p<0{\cdot}05$, ** = $p<0{\cdot}01$, *** = $p<0{\cdot}001$, ****=$p<0{\cdot}0001$.

Figure 5.

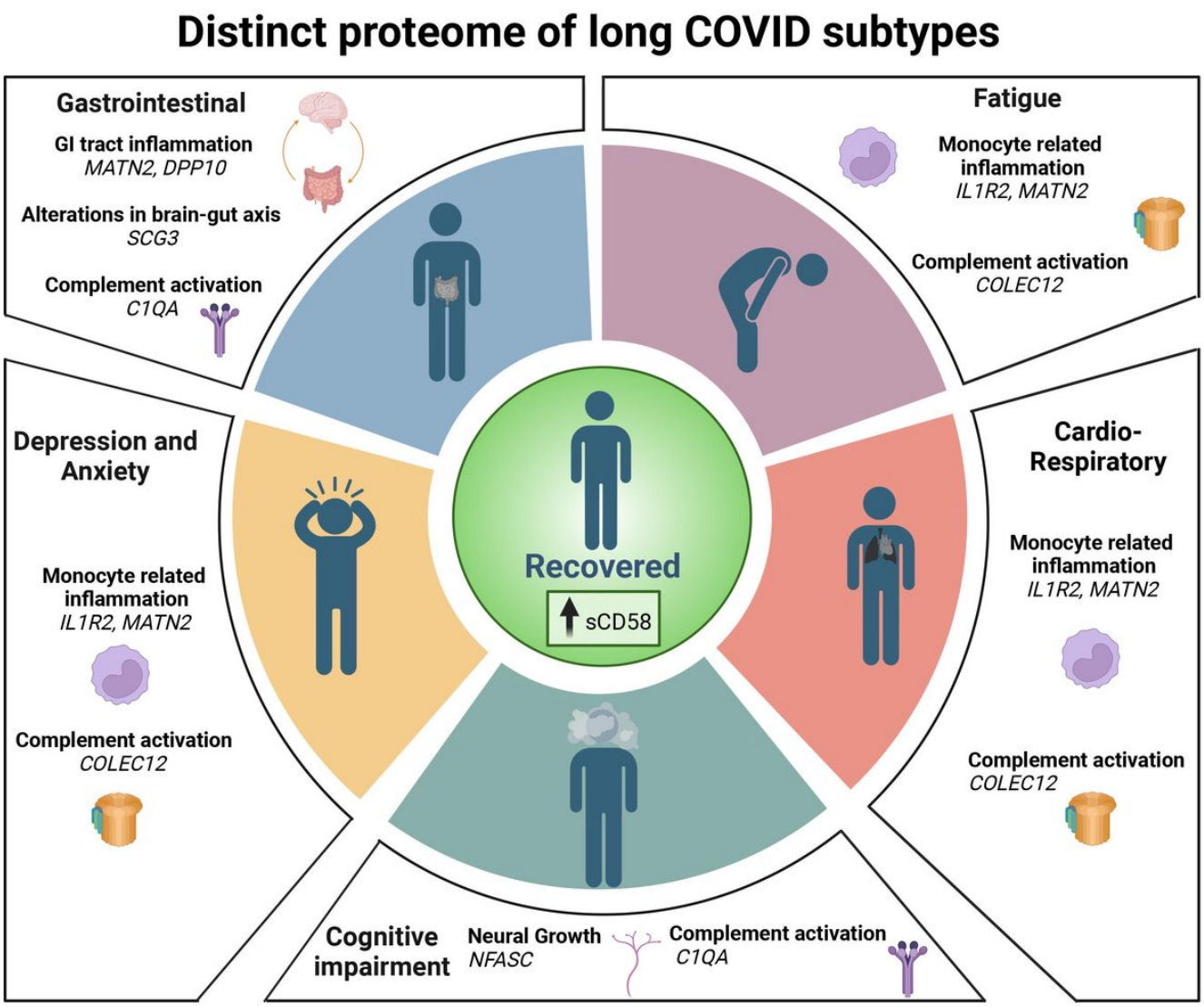

Figure 5

Graphical Abstract. Protein signatures associated with each long COVID subtype are shown. The blood proteome of 719 patients was analysed, 6 months after COVID-19 hospitalisation. For all markers shown, elevated levels were associated with each symptom outcome. Elevated sCD58 was associated with feeling recovered.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Appendix1SuppTable6NILE36351T.pdf
- supplementarymaterialsrevised3NILE36351T04oct23.docx
- FINALEXTFIG091023.pdf