SUPPLEMENTARY MATERIAL

Supplementary Methods

Inclusion criteria of DISCOVER study

Adult patients (>18yo) admitted to North Bristol NHS Trust with;

(a) typical symptoms of COVID-19 (e.g. influenza-like illness with fever and muscle pain, or respiratory illness with cough and shortness of breath) and a positive PCR result for SARS-CoV-2, using the established PHE assay in use at the time,

Or

(b) Suspected SARS-CoV-2 infection, namely presenting with (i) typical symptoms (e.g. influenza-like illness with fever and muscle pain, or respiratory illness with cough and shortness of breath); and (ii) compatible chest X-ray findings (consolidation or ground-glass shadowing); and (iii) alternative causes have been considered unlikely or excluded (e.g. heart failure, influenza).

Baseline assessment

Routine demographics were recorded including ethnicity, and presence of important comorbidities. The earliest admission National Early Warning Score (NEWS) was extracted from the clinical record. This is a numeric score (from 1-20), reflecting the degree of physiological dysfunction. Routine biochemistry and haematology results were extracted from the clinical record (C-reactive protein (CRP), neutrophils, lymphocytes, neutrophil:lymphocyte ratio), using the admission results. Chest radiography was performed on admission and radiological severity score calculated (see below).

28-day remote follow-up

All recruited patients were followed up remotely at 28-days after admission by review of hospitals notes and/or general practice records. This included 28-day mortality, hospital length of stay, readmissions, requirement for intensive care, ventilation, renal replacement therapy, and inotropes. We also recorded complications including acute renal failure, acute liver injury, venous thromboembolic events (both pulmonary emboli and deep vein thromboses), cardiac events (including myocardial infarction, myocarditis, congestive cardiac failure and arrythmias), and neurological events (cerebrovascular events, meningitis or encephalitis).

At 28 days, surviving participants were defined as having had severe disease (requirement for non-invasive ventilation (NIV), intensive care or high dependency unit admission), moderate disease (requirement for oxygen during hospital stay), or mild disease (no requirement for oxygen or enhanced care during stay).

8-12 week face-to-face outpatient follow up.

All patients who survived were invited to a follow up at a respiratory outpatient clinic (with the exception of nursing home residents or current hospital inpatients), 8-12 weeks after hospital admission. Patients were followed

up a median of 83 days (IQR 74-88 days) after hospital admission and 90 days (IQR 80-97 days) after COVID-19 symptom onset. All patients who attended this appointment had a face-to-face review with a respiratory or infectious disease clinician, chest radiograph, spirometry, exercise testing, routine bloods, routine observations (blood pressure, heart rate, temperature, pulse oximetry, respiratory rate) and HRQoL questionnaires (see details below).

Chest radiograph

Non-portable radiography equipment was used to obtain posterior-anterior (PA) projection radiographs with standard techniques at a 180-cm focus-film distance.

The radiological severity score was calculated for the baseline radiograph using the method described by Wong et al, 2020.[1] A score of 0-4 was assigned to each lung depending on the extent of involvement by consolidation or ground glass opacities. 0 = no involvement, 1 = <25%, 2 = 25 - 49%, 3= 50 - 75%, 4 = >75% involvement. The scores for each lung were summed to produce a final severity score ranging from 0-8. Radiographs were scored by one physician (respiratory or infectious diseases physician).

All follow-up chest radiographs were categorised into two groups, normal or abnormal, based on lung parenchymal, airway, pleural, hilar and mediastinal findings as reported by a consultant radiologist. In those chest x-rays demonstrating an abnormality, the lung parenchyma and airways were evaluated for the following: 1) consolidation, 2) ground-glass opacity (GGO), 3) nodular opacity, and 4) reticular opacity 5) atelectasis 6) pleural pathology, by consultant radiologists and according to standardised terminology.[2]

Spirometry

Forced expiratory volume during first second of expiration (FEV1) and forced vital capacity (FVC) were performed in accordance with ATS/ERS guidelines.[3] The MRC score, height (meters), and body weight of the patients (kilograms) were also recorded.[4] Restrictive spirometry was defined by a FEV1/FVC ratio <0.7 AND FVC <80%. [5]

Lung physiology staff wore full personal protective equipment (PPE) during testing including FFP-3 masks.

Sit to stand test (STS)

Given the importance of social distancing in clinical areas, exercise testing was assessed using the 1-min sit-to-stand test (STS) as opposed to the 6-minute walk test (6MWT). All 1-min STS tests were performed according to a standardised protocol using a standard chair (height 46–48 cm) with a flat seat and no armrests. Patients were instructed to stand completely straight from a seated position and touch the chair with their bottom when sitting,

but that they need not sit fully back on the chair. Patients were asked to complete the manoeuvre without using their hands or arms to assist movement and to perform as many repetitions as possible in 1 min. A minimum of three sit to stands were required in order for this to be recorded as an adequate test. The resting oxygen saturation was recorded via pulse oximetry, in addition to the nadir oxygen saturation during the test and up to one minute during recovery. A mild desaturation was classified as ≥ 4% but with a nadir ≥94%, a significant desaturation was classified as any desaturation with a nadir <94%.[6]

Health status questionnaires

The SF-36 is a questionnaire of 36 items, measuring eight multi-item variables; physical functioning (PF), social functioning (SF), role limitations due to physical (RP) or emotional problems (RE), mental health (MH), energy and vitality (VT), bodily pain (BP) and general perception of health (GH).[7] With regards to the measurement of mental health, the SF-36 measures general mental health status, that includes four major mental health dimensions: anxiety, depression, loss of behavioural/emotional control, and psychological well-being. There is a further single item for perception of change in health over the past year. For each variable, items are scored and transformed into a scale of 0 to 100 (best possible health status). Subsequently, a composite physical and mental composite score (PCS, MCS) are generated from each individual variant. The WEMWBS is a scale of 14 positively worded items focusing on the positive aspects of mental health and measuring psychological functioning: optimism, autonomy, agency, curiosity, clarity of thought, positive relationships and positive affect (feelings), with 5 response categories of 'none of the time' 'to 'all of the time'. Scores ranged from 14-70, with higher scores indicating greater positive mental wellbeing.[8]

Statistical analysis

Categorical variables were presented as counts with percentages. All continuous data were non-parametric and therefore presented with medians and interquartile range (IQR), unless otherwise specified. Differences between patient groups were evaluated using Mann Whitney-U and Kruskal Wallis tests for continuous data and Fisher's exact test or Chi-squared testing for categorical data. Statistical significance was taken as $p \leq 0.05$. Data were analysed using R version 4.0.0 with the packages "tidyverse" and "gtsummary".

Supplementary Table 1: Demographics and admission factors of followed-up cohort (n=110) and those who did not attend follow-up (n=34)

Characteristic	**Followed-up (n=110)**	**Did not attend follow-up* (n=34)**
Demographics		
Age (18+)	60 (IQR: 44-76)	71 (38-81)
BAME	23 (21%)	4 (12%)
Male	68 (61%)	12 (35%)
Co-morbidities		
T1DM	3 (3%)	0 (0%)
T2DM	16 (15%)	10 (29%)
Heart disease	20 (18%)	7 (21%)
Chronic Lung disease	28 (25%)	1 (3%)
Severe Liver disease	1 (1%)	4 (9%)
Severe kidney disease	7 (6%)	6 (18%)
Hypertension	27 (25%)	11 (32%)
HIV	1 (1%)	0 (0%)
Laboratory Testing		
SARS CoV-2 PCR+ve (as inpatient)	81 (73%)	25 (73%)
SARS CoV-2 Antibody +ve (at follow-up)- Abbott antibody test	89 (81%)	N/A
Admission (ED) NEWS score (IQR)	4 (2-6)	4 (2-5)
Radiographic severity score on admission chest radiograph (IQR)	2 (1-4)	2 (2-4)

BAME- Black, Asian and Minority Ethnic, DM- Diabetes mellitus, HIV- Human Immunodeficiency Virus, NEWS- National Early Warning Score.
**of 34 patients who did not attend follow up clinic, 12 were in Mild severity group, 22 in Moderate, 0 in Severe.*

Supplementary Table 2: Ongoing symptoms reported at follow-up by severity of disease

Symptom reported	**Mild** (n = 27)	**Moderate** (n = 65)	**Severe** (n = 18)
Fever	0 (0%)	1 (2%)	0 (0%)
Cough	2 (7%)	10 (15%)	1 (6%)
Breathlessness	7 (26%)	26 (40%)	10 (56%)
Anosmia	3 (11%)	6 (9.2%)	4 (22%)
Excessive Fatigue	7 (26%)	26 (40%)	10 (56%)
Myalgia	4 (15%)	14 (22%)	7 (39%)
Headache	1 (4%)	1 (2%)	0 (0%)
Chest pain	2 (7.4%)	10 (15%)	2 (11%)
Arthralgia	1 (4%)	1 (2%)	3 (16%)
Diarrhoea	0 (0%)	1 (2%)	0 (0%)
Abdominal pain	1 (2%)	1 (2%)	0 (0%)
Nausea	0 (0%)	0 (0%)	0 (0%)
Insomnia	6 (22%)	11 (17%)	9 (50%)
Any symptom	16 (59%)	49 (75%)	16 (89%)

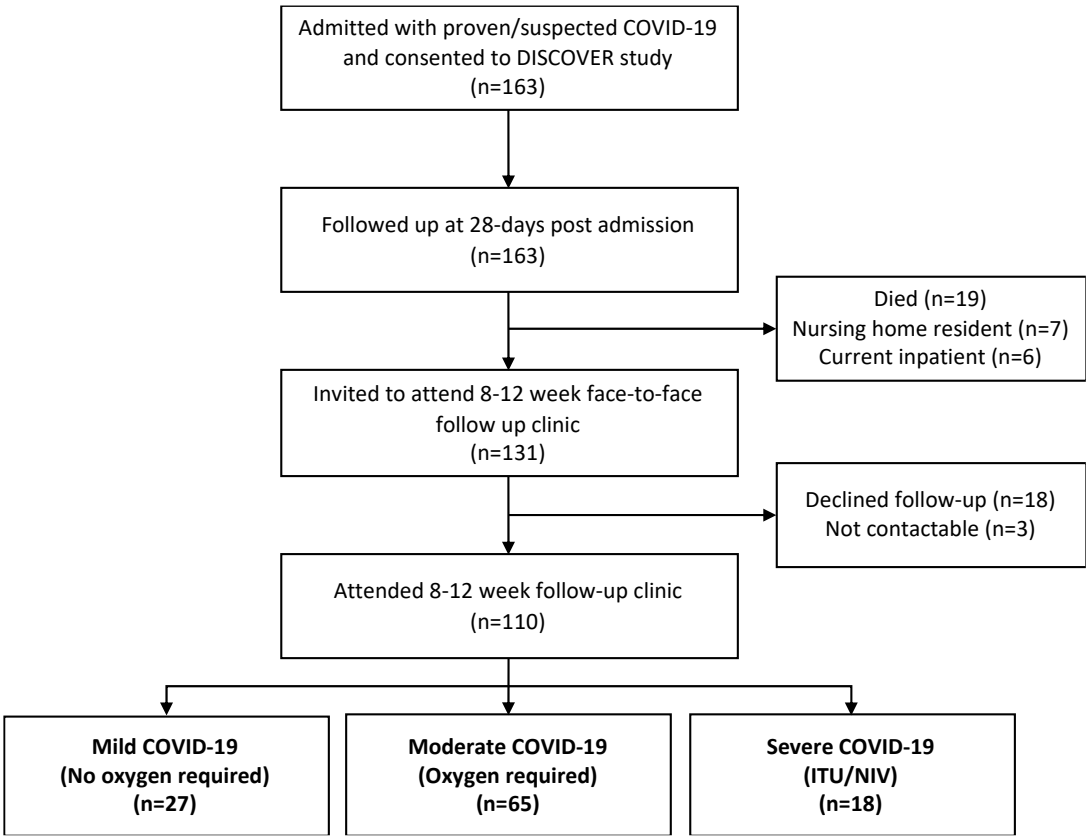
Supplementary Table 3: Spirometry and Sit-to-Stand desaturation test results

	Mild (n = 27)	**Moderate (n = 65)**	**Severe (n = 18)**	**p-value**
O2 Saturations (%)	98.0 (96.5, 99.0)	97.00 (96.0, 98.00)	97.0 (96.0, 98.0)	0.88
Nadir of O2 saturations on STS test (IQR)	96.0 (95.0, 97.0)	95.0 (93.0, 96.5)	95.0 (91.8, 96.0)	0.75
Respiratory rate (IQR)	17.0 (14.0, 18.0)	17.0 (14.2, 19.8)	17.0 (16.0, 18.0)	0.95
FVC (L) (IQR)	3.58 (3.13, 4.31)	3.52 (2.75, 4.36)	3.65 (2.55, 4.14)	0.70
FVC (% predicted) (IQR)	97 (90, 105)	91 (78, 100)	89 (76, 98)	0.05
FEV1 (L) (IQR)	2.97 (2.56, 3.42)	2.71 (2.12, 3.49)	2.54 (1.88, 3.23)	0.50
FEV1 (% predicted) (IQR)	94 (82, 101)	90 (78, 100)	89 (73, 101)	0.30
Spirometry not performed*	0/27	5/65	0/18	N/A
Restrictive pattern spirometry (%)	0 (0%)	8 (12%)	3 (17%)	0.03
Severe desaturation on STS test (%)	0 (0%)	10 (15%)	5 (28%)	0.02
*STS- 1 minute Sit to Stand desaturation test, FVC- Forced Vital Capacity, FEV1- Forced expiratory volume during first second of expiration. *Spirometry not performed for clinical reason (n=3) and patient declined (n=2).*				

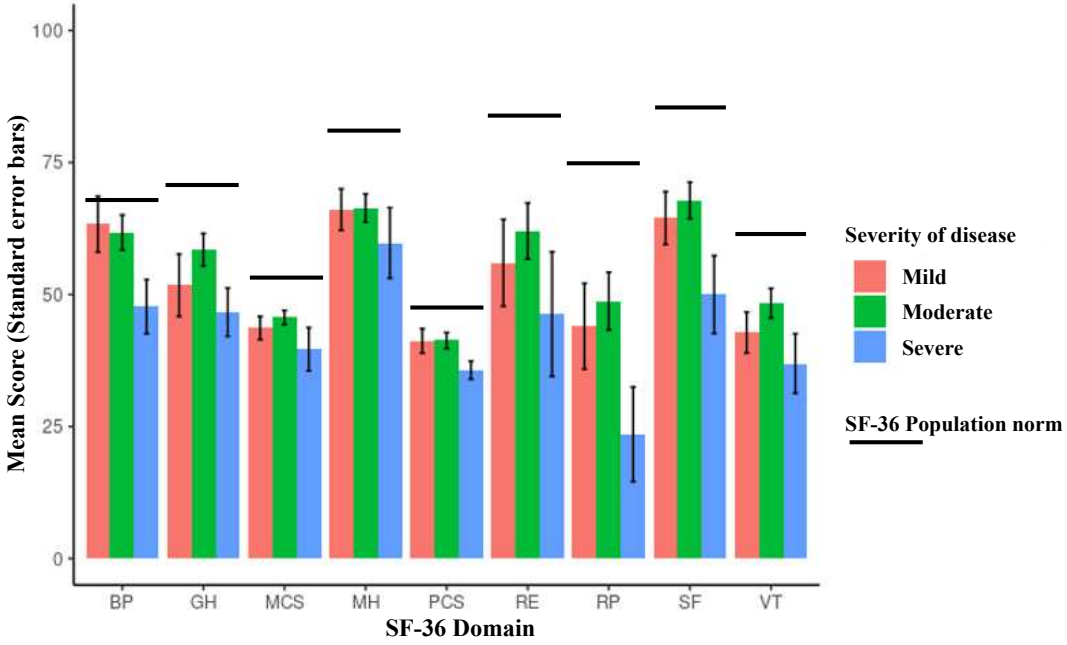
Supplementary Table 4: Full SF-36 results by severity of disease

Characteristic*	**Mild** (n = 27)	**Moderate** (n = 65)	**Severe** (n = 18)
Physical Function	66 (27)	65 (30)	56 (23)
Role – physical	43 (41)	50 (43)	24 (38)
Bodily pain	64 (28)	62 (27)	46 (21)
General Health	53 (30)	58 (25)	47 (20)
Vitality	43 (20)	49 (22)	36 (24)
Social Functioning	65 (25)	69 (28)	49 (32)
Role - emotional	57 (42)	64 (42)	43 (50)
Mental health	68 (19)	67 (21)	58 (28)
Physical composite score	41 (12)	41 (12)	36 (7)
Mental composite score	45 (11)	46 (11)	40 (17)
*Statistics presented: mean (SD)			

Supplementary Figure 1: CONSORT diagram



Supplementary Figure 2: SF-36 results, mean and standard error, with age-matched population norm means (BP- Bodily pain, GH- General Health, MCS- Mental Health Component Summary, MH-Mental Health, PCS- Physical Component Summary, RE- Emotional Role Functioning, RP-Physical Role Functioning, SF-Social Role Functioning, VT-Vitality).



REFERENCES

1. Wong HYF, Lam HYS, Fong AH, et al. Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19. *Radiology* 2020;296(2):E72-E78. doi: 10.1148/radiol.2020201160 [published Online First: 2020/03/29]
2. Hansell DM, Bankier AA, MacMahon H, et al. Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008;246(3):697-722. doi: 10.1148/radiol.2462070712 [published Online First: 2008/01/16]
3. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. *Am J Respir Crit Care Med* 2019;200(8):e70-e88. doi: 10.1164/rccm.201908-1590ST [published Online First: 2019/10/16]
4. Bestall JC, Paul EA, Garrod R, et al. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999;54(7):581-6. doi: 10.1136/thx.54.7.581 [published Online First: 1999/06/22]
5. Aaron SD, Dales RE, Cardinal P. How accurate is spirometry at predicting restrictive pulmonary impairment? *Chest* 1999;115(3):869-73. doi: 10.1378/chest.115.3.869 [published Online First: 1999/03/20]
6. Briand J, Behal H, Chenivesse C, et al. The 1-minute sit-to-stand test to detect exercise-induced oxygen desaturation in patients with interstitial lung disease. *Ther Adv Respir Dis* 2018;12:1753466618793028. doi: 10.1177/1753466618793028 [published Online First: 2018/08/10]
7. Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: Scoping review. *SAGE Open Med* 2016;4:2050312116671725. doi: 10.1177/2050312116671725 [published Online First: 2016/10/21]
8. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes* 2007;5:63. doi: 10.1186/1477-7525-5-63 [published Online First: 2007/11/29]