ONLINE SUPPLEMENT

Cluster Analysis to Identify Long COVID Phenotypes Using ^{129}Xe Magnetic Resonance Imaging: A Multi-centre Evaluation

SUPPLEMENTARY METHODS

Study Participants and Design

Participants aged ≥19-years provided written informed-consent to ethics board-approved protocols between July 2021 and February 2023 across the three centres. Tests were performed by real-time reverse transcription polymerase chain reaction or lateral flow test depending on local guidelines at the time of testing. Self-reported COVID positive test dates and subsequently MRI study enrollment dates at each were:

- Site1 University of British Columbia: infections August 2020 – September 2021, MRI enrollment January 2022 – February 2023
- Site2 University of Kansas Medical Center: infections September 2020 – November 2021, MRI enrollment September 2021 – November 2022
- Site3 Duke University: infections September 2020 – August 2021, MRI enrollment July 2021 – June 2022

MRI Acquisition and Analysis

Total inhaled doses were 1.0L (Site1) or 20% of the participants FVC rounded to the nearest 200mL up to a maximum of 1.25L (Site2, Site3), and participants were instructed to inhale the gas mixture from a Tedlar® bag (Jensen Inert Products, Coral Springs, USA) from functional residual capacity for imaging under breath-hold conditions. Up to 1.0L ^{129}Xe gas was polarised using 9820 hyperpolarisers (Polarean Imaging Inc., Durham, USA) and administered ^{129}Xe volumes were expressed as the dose equivalent volume (DEV), which represents an equivalent ^{129}Xe volume that is 100% isotopically enriched and 100% hyperpolarised [1] to enable comparisons between sites.

Participants first performed an in vivo calibration scan [2] using 100-200mL ^{129}Xe diluted to the total dose with N_2 to determine participant-specific parameters including xenon resonance frequency, transmitter power, TE90 (echo time at which the dissolved ^{129}Xe red blood cell [RBC] and membrane [Mem] signals are 90° out of phase) and RBC/Mem ratio. Gas exchange imaging was then performed using single-breath 1-point Dixon methods. Site1 employed a standard 1-point Dixon method with identical resolution for gas and dissolved phase according to current recommended guidelines [2,3]. Site2 employed an interleaved spiral/radial approach that encodes a high-resolution ventilation image alongside standard resolution dissolved images in a single breath-hold [4], and Site3 employed a flip angle/repetition time equivalence accelerated 1-point Dixon method [5]. While non-harmonised protocols were used because of the post-hoc aggregation of multi-centre data, in all cases the $TR_{90,equiv}$ was 249 ms, enabling comparison of gas exchanges values [5,6]. Anatomical ^{1}H imaging was performed directly following ^{129}Xe gas exchange; Site1 and Site3 acquired ^{1}H images in a separate breath-hold, and Site2 performed ^{1}H imaging during the same breath-hold as ^{129}Xe gas exchange to facilitate image quantification.

Image segmentation was first performed with an automated machine learning algorithm [7], and resulting thoracic cavity masks were manually edited by trained observers as needed. Membrane and RBC dissolved images were divided on a voxel-wise basis by the gas-phase image intensities to create normalised images of membrane uptake and RBC transfer, respectively, and normalised voxel intensities were averaged to get the mean signal value [8].

Long COVID Cluster Internal Validation

Internal cluster validation was evaluated using the Davies-Bouldin index [9], silhouette width [10] and Dunn's index [11], which are measures of the average similarity between clusters, average distance between clusters, and the ratio of the smallest distance between observations in

different clusters to largest cluster diameter, respectively. A minimal Davies-Bouldin index and maximal silhouette width and Dunn's index indicate better quality clustering.

SUPPLEMENTARY RESULTS

The number of participants in each study group significantly differed by study site ($p<0.001$), driven by the lack of recovered participants that were not recruited by design at Site3. Tables E1-E3 show study measurements by site for never-COVID, recovered and long COVID groups, respectively. Most noteworthy, MRI at Site1 was performed significantly later after COVID diagnosis than Site2 for fully recovered (479±231vs. 237±185 days, $p=0.008$) and long COVID participants (534±213 vs. 356±182 days, $p=0.002$), and significantly later than Site3 for long COVID participants (vs. 242±187 days, $p<0.001$). Long COVID participants at Site3 had generally reduced pulmonary function test measurements compared with the other two sites ($p<0.05$ vs. at least Site1 or Site2). Although significantly different across sites for some groups, SNR was adequate for analysis in all cases and was not significantly different between combined study groups, supporting aggregation of the multi-center data for clustering. We note that membrane and RBC SNR are dependent on the ^{129}Xe transfer efficiency (Mem/Gas, RBC/Gas, RBC/Mem) which are increased or reduced in some patients, thus membrane and RBC SNR are not directly interpretable.

Participants from each site were distributed throughout all clusters (Table E4). There were no differences in prevalence of specific comorbidities across the clusters, nor the total number of comorbidities by participant ($p=0.2$, Table E5). The median number of comorbidities was significantly different across the clusters ($p=0.04$, Table E5), with Cluster3 and Cluster4 trending the greatest number of comorbidities ($p=0.055$ and $p=0.054$, respectively vs Cluster1). Tables

E6-E17 show full cluster-wise comparisons with post-hoc Bonferroni corrected p-values. Because age has been shown to be significantly related to ^{129}Xe MRI RBC/Mem [12,13], we evaluated age relationships to separate potential contributions of age versus long COVID to differences in RBC/Mem across groups and clusters. As shown in Figure E3, RBC/Mem was significantly negatively correlated with age in never-COVID (ρ=-0.68, $p<0.001$), recovered (ρ=-0.35, $p=0.04$) and long COVID (ρ=-0.42, $p<0.001$). The age ranges were overlapping across the groups though, suggesting that age-RBC/Mem relationships did not largely contribute to cluster separation. BMI was not significantly related to RBC/Mem in the COVID-negative nor recovered groups (both $p=0.7$). For all long COVID participants, BMI was significantly but weakly correlated with RBC/Mem (ρ=-0.31, $p=0.007$), although likely driven by two participants with very obese BMIs (49, 60 kg/m2). The BMI ranges overlapped across all groups, suggesting that BMI differences across clusters are true physiological characteristics. BMI was highest in Cluster4, and could have could have been a risk factor for these patients for more severe SARS-CoV-2 infections (75% were hospitalised during their acute infection) as well as development of long COVID.

Table E1. COVID-negative Participant Demographics and Study Measurements by Site.

Parameter ± SD	ALL (n=28)	Site1 (n=11)[1]	Site2 (n=7)[2]	Site3 (n=10)[3]	p-value*
Age years	40 ± 16	44 ± 18	33 ± 13	41 ± 16	0.3
Female n (%)	16 (57)	8 (73)	3 (43)	5 (50)	0.4
Caucasian n (%)	22 (79)	7 (64)	7 (100)	8 (80)	0.5
BMI kg/m^2	27 ±4	25 ± 3	27 ± 6	28 ± 3	0.2
Dx days n	-	-	-	-	-
FEV_1 $\%_{pred}$	105 ± 15	106 ± 14	104 ± 15	103 ± 17	0.9
FVC $\%_{pred}$	107 ± 13	109 ± 13	105 ± 9	107 ± 16	0.9
FEV_1/FVC %	81 ± 8	81 ± 9	82 ± 6	80 ± 7	0.8
RV $\%_{pred}$	108 ± 25	113 ± 17	-	93 ± 40	0.2
TLC $\%_{pred}$	99 ± 13	102 ± 11	-	96 ± 14	0.4
RV/TLC $\%_{pred}$	107 ± 19	112 ± 15	-	94 ± 23	0.09
DL_{CO} $\%_{pred}$	103 ± 14	105 ± 16	-	99 ± 12	0.3
VDP %	0 ± 1	0 ± 0	0 ± 1	1 ± 1†	**0.04**
Mem/Gas	0.71 ± 0.15	0.69 ± 0.16	0.76 ± 0.18	0.70 ± 0.11	0.6
RBC/Mem	0.47 ± 0.13	0.39 ± 0.11	0.53 ± 0.12	0.51 ±0.13	**0.046**
RBC/Gas	0.35 ± 0.12	0.29 ± 0.11	0.40 ± 0.11	0.38 ± 0.13	0.1
DEV mL	190 ± 36	191 ± 21	204 ± 53	178 ± 34	0.4
Gas SNR	21 ± 6	22 ± 3	25 ± 7	16 ± 4†‡	**0.003**
Mem SNR	19 ± 8	19 ± 5	28 ± 7†	14 ± 6‡	**<0.001**
RBC SNR	9 ± 5	8 ± 3	15 ± 4†	7 ± 4‡	**<0.001**

BMI=body mass index; DEV=dose equivalent volume; Dx days=days between COVID diagnosis and MRI; DL_{CO}=diffusion capacity of the lungs for carbon monoxide; FEV_1=forced expiratory volume in one second; FVC=forced vital capacity; Mem=interstitial membrane; RBC=red blood cells; RV=residual volume; SNR=signal-to-noise ratio; TLC=total lung capacity; VDP=ventilation defect percent; $\%_{pred}$=percent predicted.

*Using one-way ANOVA for parametric variables or Kruskal-Wallis tests for non-parametric variables.

[1]All participants at Site1 have all measurements; [2]Site2 n=6 for FEV_1, FVC, FEV_1/FVC, n=0 for TLC and DL_{CO}; [3]Site3 n=7 for TLC.

†Significantly different from Site1 ($p<0.05$); ‡Significantly different from Site2 ($p<0.05$)

Table E2. Recovered Participant Demographics and Study Measurements by Site.

Parameter ± SD	ALL (n=35)	Site1 (n=25)[1]	Site2 (n=9)[2]	Site3 (n=0)[3]	p-value*
Age years	43 ± 14	45 ± 14	37 ± 14	-	0.1
Female n (%)	20 (57)	14 (56)	6 (67)	-	0.7
Caucasian n (%)	28 (80)	19 (76)	9 (100)		0.8
BMI kg/m^2	27 ± 5	27 ± 5	27 ± 4	-	0.8
Dx days n	417 ± 243	479 ± 231	237 ± 185	-	**0.008**
FEV_1 $\%_{pred}$	107 ± 20	107 ± 20	99	-	-
FVC $\%_{pred}$	110 ± 16	111 ± 16	93	-	-
FEV_1/FVC %	79 ± 9	79 ± 9	87	-	-
RV $\%_{pred}$	112 ± 23	113 ± 21	65	-	-
TLC $\%_{pred}$	102 ± 11	102 ± 10	79	-	-
RV/TLC $\%_{pred}$	112 ± 24	113 ± 24	83	-	-
DL_{CO} $\%_{pred}$	108 ± 10	108 ± 10	104	-	-
VDP %	0 ± 1	0 ± 1	1 ± 1	-	0.5
Mem/Gas	0.73 ± 0.18	0.74 ± 0.15	0.72 ± 0.25	-	0.7
RBC/Mem	0.44 ± 0.11	0.44 ± 0.10	0.42 ± 0.14	-	0.6
RBC/Gas	0.33 ± 0.09	0.34 ± 0.09	0.30 ± 0.09	-	0.3
DEV mL	197 ± 47	211 ± 32	155 ± 60	-	**0.001**
Gas SNR	22 ± 6	23 ± 5	20 ± 8	-	0.2
Mem SNR	22 ± 8	20 ± 5	25 ± 14	-	0.2
RBC SNR	9 ± 3	9 ± 3	10 ± 4	-	0.9

BMI=body mass index; DEV=dose equivalent volume; Dx days=days between COVID diagnosis and MRI; DL_{CO}=diffusion capacity of the lungs for carbon monoxide; FEV_1=forced expiratory volume in one second; FVC=forced vital capacity; Mem=interstitial membrane; RBC=red blood cells; RV=residual volume; SNR=signal-to-noise ratio; TLC=total lung capacity; VDP=ventilation defect percent; $\%_{pred}$=percent predicted.

*Using independent samples t-tests for parametric variables or Mann-Whitney U-tests for non-parametric variables.

[1]All participants at Site1 have all measurements; [2]Site2 n=1 for all pulmonary function tests; [3]Site3 did not recruit fully recovered COVID participants by design.

Table E3. Long COVID Participant Demographics and Study Measurements by Site.

Parameter ± SD	ALL (n=73)	Site1 (n=32)[1]	Site2 (n=28)[2]	Site3 (n=13)[3]	p-value*
Age years	49 ± 13	48 ± 11	50 ± 13	50 ± 12	0.8
Female n (%)	44 (60)	18 (56)	20 (71)	6 (46)	0.3
Caucasian n (%)	62 (85)	22 (69)	27 (96)	13 (100)	0.4
BMI kg/m^2	29 ± 6	28 ± 6	29 ± 7	29 ± 4	0.6
Dx days n	414 ± 226	534 ± 213	356 ± 182†	242 ± 187†‡	**<0.001**
FEV_1 $\%_{pred}$	99 ± 18	104 ± 18	98 ± 17	91 ± 19†	0.08
FVC $\%_{pred}$	102 ± 20	109 ± 19	99 ± 16	93 ± 21†	**0.02**
FEV_1/FVC %	79 ± 6	78 ± 6	81 ± 6†	80 ± 6‡	0.2
RV $\%_{pred}$	106 ± 30	117 ± 26	94 ± 18	85 ± 36	**0.001**
TLC $\%_{pred}$	94 ± 18	102 ± 17	87 ± 13	84 ± 19†	**0.001**
RV/TLC $\%_{pred}$	112 ± 22	116 ± 19	110 ± 16	103 ± 35	0.4
DL_{CO} $\%_{pred}$	99 ± 24	104 ± 20	102 ± 19	80 ± 32‡	0.06
VDP %	1 ± 2	0 ± 0	1 ± 1	2 ± 4†	**<0.001**
Mem/Gas	0.77 ± 0.24	0.72 ± 0.19	0.75 ± 0.23	0. 93 ± 0.33	0.1
RBC/Mem	0.40 ± 0.12	0.43 ± 0.12	0.37 ± 0.10	0.39 ± 0.14	0.2
RBC/Gas	0.31 ± 0.09	0.31 ± 0.08	0.29 ± 0.10	0.36 ± 0.08	0.053
DEV mL	186 ± 42	203 ± 41	172 ± 38†	178 ± 40	**0.01**
Gas SNR	22 ± 7	23 ± 4	22 ± 8	20 ± 8	0.5
Mem SNR	22 ± 10	20 ± 7	27 ± 10†	15 ± 9‡	**<0.001**
RBC SNR	9 ± 4	8 ± 3	10 ± 5	5 ± 2†‡	**<0.001**

BMI=body mass index; DEV=dose equivalent volume; Dx days=days between COVID diagnosis and MRI; DL_{CO}=diffusion capacity of the lungs for carbon monoxide; FEV_1=forced expiratory volume in one second; FVC=forced vital capacity; Mem=interstitial membrane; RBC=red blood cells; RV=residual volume; SNR=signal-to-noise ratio; TLC=total lung capacity; VDP=ventilation defect percent; $\%_{pred}$=percent predicted.

*Using one-way ANOVA for parametric variables or Kruskal-Wallis tests for non-parametric variables.

[1]All participants at Site1 have all measurements; [2]Site2 n=22 for FEV_1, FVC, FEV_1/FVC, n=18 for TLC, n=17 for DL_{CO}; [3]Site3 n=10 for TLC, n=12 for DL_{CO}.

†Significantly different from Site1 ($p<0.05$); ‡Significantly different from Site2 ($p<0.05$

Table E4. Number of Participants in Each Cluster by Site.

	Cluster1 (n=24)	Cluster2 (n=22)	Cluster3 (n=19)	Cluster4 (n=8)
Site1 n (%)	15 (62)	6 (27)	8 (42)	3 (38)
Site2 n (%)	6 (25)	14 (64)	7 (37)	1 (12)
Site3 n (%)	3 (13)	2 (9)	4 (21)	4 (50)

Table E5. Interventions during Acute COVID Hospitalisation in Each Cluster.

	ALL (n=73)	**Cluster1 (n=24)**	**Cluster2 (n=22)**	**Cluster3 (n=19)**	**Cluster4 (n=8)**	**p-value***
COVID Hosp n (%)	18 (25)	3 (13)	7 (32)**	3 (16)	5 (63)	**0.03**
Hosp Days mean ± SD	13 ± 14	5 ± 1	9 ± 10	7 ± 3	28 ± 18	**0.03**
Hosp Days median [IQR]	6 [4,22]	5 [5,5]	5 [4,7]	8 [6,9]	24 [21,37]	0.06
Oxygen	11	3 (100)	2 (29)	1 (33)	5 (100)	**0.005**
ICU	8	0 (0)	3 (43)	1 (33)	4 (80)	**0.001**
Ventilation	3	0 (0)	0 (0)	0 (0)	3 (60)	**<0.001**
Type of Ventilation						
Invasive	3	-	-	-	3 (100)	-
Non-invasive	0	-	-	-	0 (0)	-
Medical Therapies						
Steroids	14 (78)	3 (100)	5 (71)	3 (100)	3 (60)	**0.045**
Monoclonal Antibodies	1 (6)	0 (0)	0 (0)	1 (33)	0 (0)	-
Antivirals	6 (33)	2 (67)	1 (14)	1 (33)	2 (40)	0.057
Antibiotics	4 (22)	1 (33)	2 (29)	0 (0)	1 (20)	0.08
Anticoagulants	1 (6)	0 (0)	1 (14)	0 (0)	0 (0)	0.05
Convalescent Plasma	1 (6)	0 (0)	1 (14)	0 (0)	0 (0)	0.05
Bronchodilators	1 (6)	0 (0)	1 (14)	0 (0)	0 (0)	0.05
Others	2 (11)	1 (33)	1 (14)	0 (0)	0 (0)	0.05

All values presented as number of participants n (%) unless otherwise stated, where the fraction in brackets represents the fraction of hospitalised patients within each cluster.
Steroids including dexamethasone, prednisone and/or not specified;
Monoclonal antibodies including tocilizumab;
Antivirals including remdesivir;
Antibiotics including piperacillin/tazobactam (Zosyn), clindamycin or not specified;
Anticoagulants including enoxaparin (Lovenox);
Bronchodilators including albuterol;
Others including potassium, sodium, sulfasalazine (disease-modifying anti-rheumatic drug), colchicine, primaquine (anti-malarial);
*Using one-way ANOVA for parametric variables, Kruskal-Wallis tests for non-parametric variables, or Fisher's exact test for categorial variables.
**COVID in-patient medical therapy information available for 6 participants.

Table E6. Number of Comorbidities by Participant in Each Cluster.

	ALL (n=73)	Cluster1 (n=24)	Cluster2 (n=22)	Cluster3 (n=19)	Cluster4 (n=8)	p-value*
Median	1	0.5	1	1	1.5	**0.04**
Frequencies						0.2
0	23	12	4	4	3	-
1	29	10	11	7	1	-
2	14	2	4	5	3	-
3	4	0	1	2	1	-
4	3	0	2	1	0	-

*Using one-way Kruskal-Wallis tests for non-parametric variables or Fisher's exact test for categorial variables.

Table E7. Pre-Existing Respiratory Disease and Respiratory Medications in Each Cluster.

	ALL (n=73)	Cluster1 (n=24)	Cluster2 (n=22)	Cluster3 (n=19)	Cluster4 (n=8)	p-value*
Asthma n (%)	20 (27)	5 (21)	8 (36)	6 (32)**	1 (13)	0.5
COPD n (%)	2 (3)	0 (0)	1 (5)	1 (5)	0 (0)	0.6
Respiratory Medications						
SABA	12 (55)	3 (6)	5 (56)	4 (57)	0	0.9
SAMA	0 (0)	0 (0)	0 (0)	0 (0)	0	-
SABA/SAMA	1 (5)	0 (0)	0 (0)	1 (14)***	0	0.6
LABA	1 (5)	1 (20)	0 (0)	0 (0)	0	0.3
LAMA	2 (9)	0 (0)	0 (0)	2 (29)***	0	0.2
LABA/LAMA	0 (5)	0 (0)	0 (0)	0 (0)	0	-
ICS	4 (18)	1 (20)	2 (22)	1 (14)	0	1.0
ICS/LABA	10 (45)	2 (40)	4 (44)***	4 (57)	0	0.8
ICS/LABA/LAMA	0 (0)	0 (0)	0 (0)	0 (0)	0	-
LTRA	1 (5)	0 (0)	1 (11)	0 (0)	0	1.0
Monoclonal Antibodies	1 (5)	0 (0)	0 (0)	1 (14)	0	0.6
OCS	1 (5)	0 (0)	0 (0)	0 (0)	1 (100)	0.05
Antibiotics	1 (5)	0 (0)	0 (0)	0 (0)	1 (100)	0.05

All values presented as number of participants n (%), where the fraction in brackets represents the fraction of patients with respiratory disease (asthma or COPD) within each cluster.

SABA=short-acting β-agonist; SAMA=short-acting muscarinic antagonist; LABA=long-acting β-agonist; LAMA=long-acting muscarinic antagonist; ICS=inhaled corticosteroid; LTRA=leukotriene receptor antagonist; OCS=oral corticosteroids.

*Using Fisher's exact test for categorical variables.

**Respiratory medications available for 5 participants with asthma.

***Including participant with COPD.

Table E8. **Groupwise Comparisons for VDP.**

ANOVA p=0.003	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	1.0	-	-	-	-
Cluster2	0.09	**0.02**	**0.005**	-	-	-
Cluster3	1.0	1.0	1.0	0.9	-	-
Cluster4	0.9	0.6	0.2	1.0	1.0	-

VDP=ventilation defect percent derived from ^{129}Xe MRI.

Table E9. **Groupwise Comparisons for Mem/Gas.**

ANOVA p<0.001	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	0.3	0.07	-	-	-	-
Cluster2	1.0	0.3	1.0	-	-	-
Cluster3	**<0.001**	**<0.001**	**<0.001**	**<0.001**	-	-
Cluster4	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	-

Mem/Gas=ratio of interstitial membrane to gas signal derived from ^{129}Xe MRI.

Table E10. **Groupwise Comparisons for RBC/Mem.**

ANOVA p<0.001	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	0.08	-	-	-	-
Cluster2	**<0.001**	**0.002**	**<0.001**	-	-	-
Cluster3	1.0	1.0	0.1	**0.03**	-	-
Cluster4	**<0.001**	**<0.001**	**<0.001**	0.08	**<0.001**	-

RBC/Mem=ratio of red blood cell to interstitial membrane signal derived from ^{129}Xe MRI.

Table E11. **Groupwise Comparisons for RBC/Gas.**

ANOVA p<0.001	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	1.0	-	-	-	-
Cluster2	**<0.001**	**<0.001**	**<0.001**	-	-	-
Cluster3	0.5	**0.02**	0.052	**<0.001**	-	-
Cluster4	1.0	1.0	1.0	0.7	**0.02**	-

RBC/Gas=ratio of red blood cell to gas signal derived from ^{129}Xe MRI.

Table E12. **Groupwise Comparisons for Age.**

ANOVA p<0.001	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	1.0	-	-	-	-
Cluster2	**<0.001**	**0.007**	**0.02**	-	-	-
Cluster3	0.4	1.0	1.0	1.0	-	-
Cluster4	0.4	1.0	1.0	1.0	1.0	-

Table E13. **Groupwise Comparisons for BMI.**

ANOVA p=0.004	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	1.0	-	-	-	-
Cluster2	1.0	1.0	0.8	-	-	-
Cluster3	0.2	0.1	0.08	1.0	-	-
Cluster4	0.07	0.051	**0.04**	0.6	1.0	-

BMI=body mass index.

Table E14. **Groupwise Comparisons for FEV_1.**

ANOVA p=0.009	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	0.9	-	-	-	-
Cluster2	1.0	1.0	1.0	-	-	-
Cluster3	1.0	1.0	1.0	1.0	-	-
Cluster4	**0.03**	**0.003**	0.2	0.2	0.06	-

FEV_1=forced expiratory volume in one second.

Table E15. **Groupwise Comparisons for FVC.**

ANOVA p<0.001	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	1.0	-	-	-	-
Cluster2	1.0	1.0	1.0	-	-	-
Cluster3	1.0	1.0	1.0	1.0	-	-
Cluster4	**0.001**	**<0.001**	**0.001**	**0.02**	**0.007**	-

FVC=forced vital capacity.

Table E16. **Groupwise Comparisons for RV.**

ANOVA p=0.008	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	1.0	-	-	-	-
Cluster2	1.0	1.0	1.0	1.0	-	-
Cluster3	1.0	1.0	1.0	1.0	-	-
Cluster4	0.054	**0.02**	**0.002**	0.1	0.1	-

RV=residual volume.

Table E17. **Groupwise Comparisons for TLC.**

ANOVA p<0.001	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	1.0	-	-	-	-
Cluster2	1.0	1.0	1.0	1.0	-	-
Cluster3	1.0	1.0	1.0	1.0	-	-
Cluster4	**<0.001**	**<0.001**	**<0.001**	**0.007**	**0.008**	-

TLC=total lung capacity.

Table E18. **Groupwise Comparisons for DL_{CO}.**

ANOVA p<0.001	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	1.0	1.0	-	-	-	-
Cluster2	0.9	**0.050**	**0.04**	-	-	-
Cluster3	1.0	1.0	1.0	0.1	-	-
Cluster4	**<0.001**	**0.001**	**<0.001**	**<0.001**	**<0.001**	-

DL_{CO}=diffusion capacity of the lungs for carbon monoxide.

Table E19. **Groupwise Comparisons for SGRQ.**

ANOVA p<0.001	Never	Recovered	Cluster1	Cluster2	Cluster3	Cluster4
Negative	-	-	-	-	-	-
Recovered	1.0	-	-	-	-	-
Cluster1	**<0.001**	**<0.001**	-	-	-	-
Cluster2	**<0.001**	**<0.001**	1.0	-	-	-
Cluster3	**<0.001**	**<0.001**	1.0	1.0	-	-
Cluster4	**<0.001**	**<0.001**	**0.03**	0.3	0.3	-

SGRQ=St. George's Respiratory Questionnaire.

Figure E1. ^{129}Xe MRI Measurements by Group.

Plots by group for MRI (A) ventilation defect percent (VDP), (B) interstitial membrane (Mem)/Gas ratio or membrane uptake, (C) red blood cell (RBC)/Mem ratio or RBC transfer, and (D) RBC/Gas ratio, highlight spread and therefore heterogeneity of MRI measurements in the total long COVID group. †Significantly different from COVID-negative ($p<0.05$)

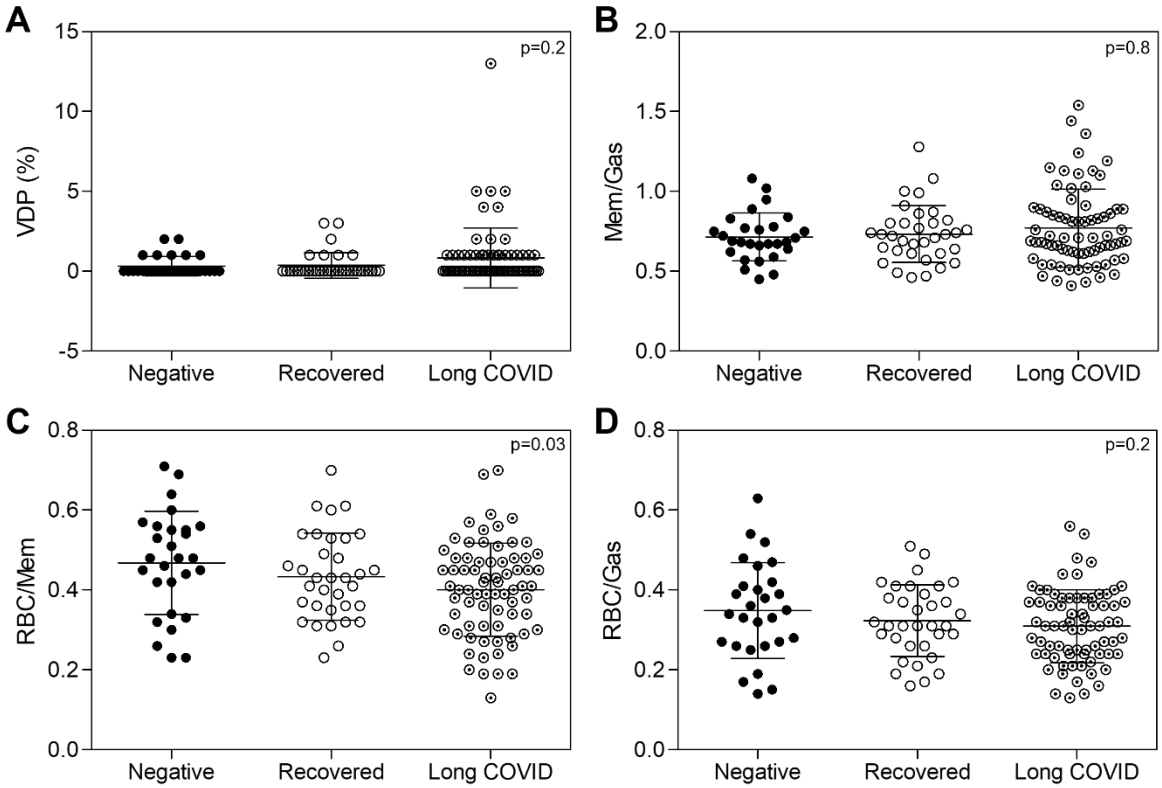


Figure E2. ^{129}Xe MRI and CT Across Long COVID Clusters.

^{129}Xe MR gas, membrane (Mem) and red blood cell (RBC) and computed tomography (CT) images for representative participants in each cluster as well as COVID-negative and recovered participants. All images shown from Site1 with a consistent CT protocol. Cluster1 participant had normal MRI findings and no CT abnormalities. Cluster2 participant had reduced MRI RBC/Mem despite no CT abnormalities. Cluster3 participant exhibited mildly elevated Mem/Gas with some reticulation on CT, whereas Cluster4 participant exhibited markedly elevated MRI Mem/Gas with ground glass, reticulation and honeycombing on CT.

D=defect (no signal); L=low-intensity signal; H=high-intensity signal.

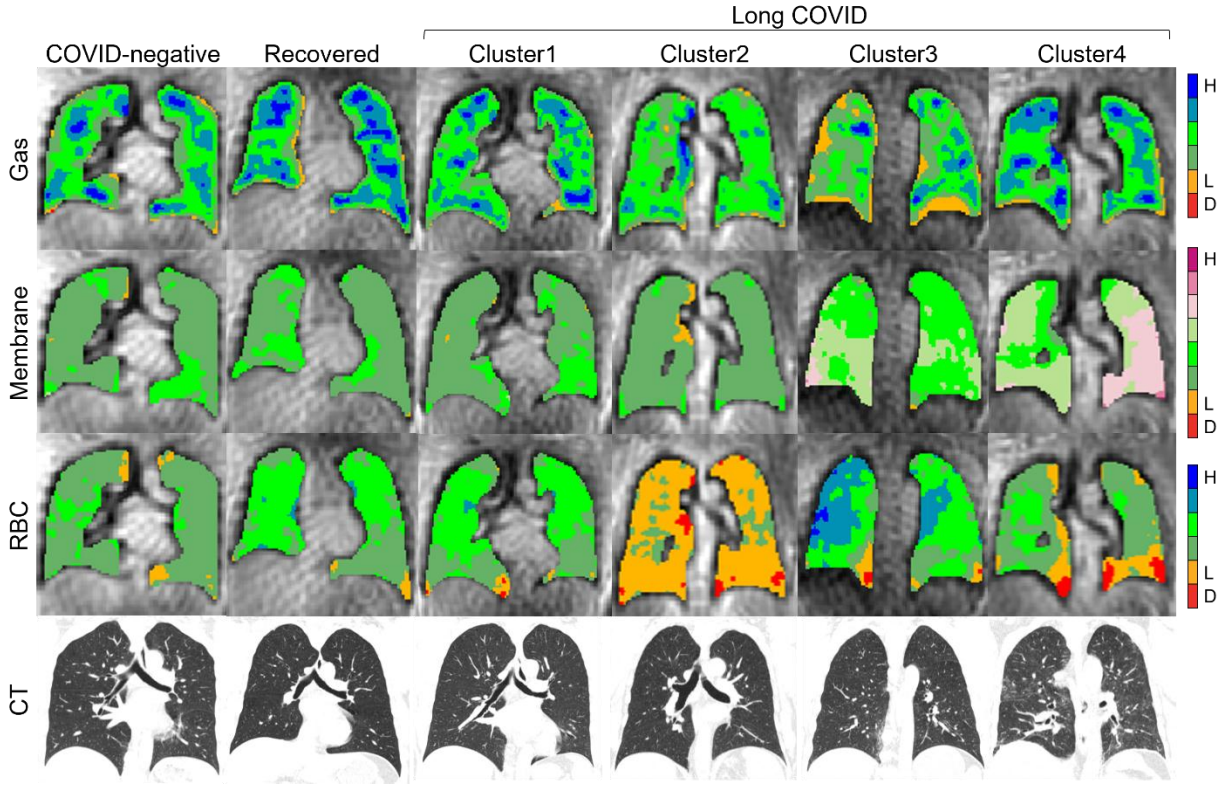
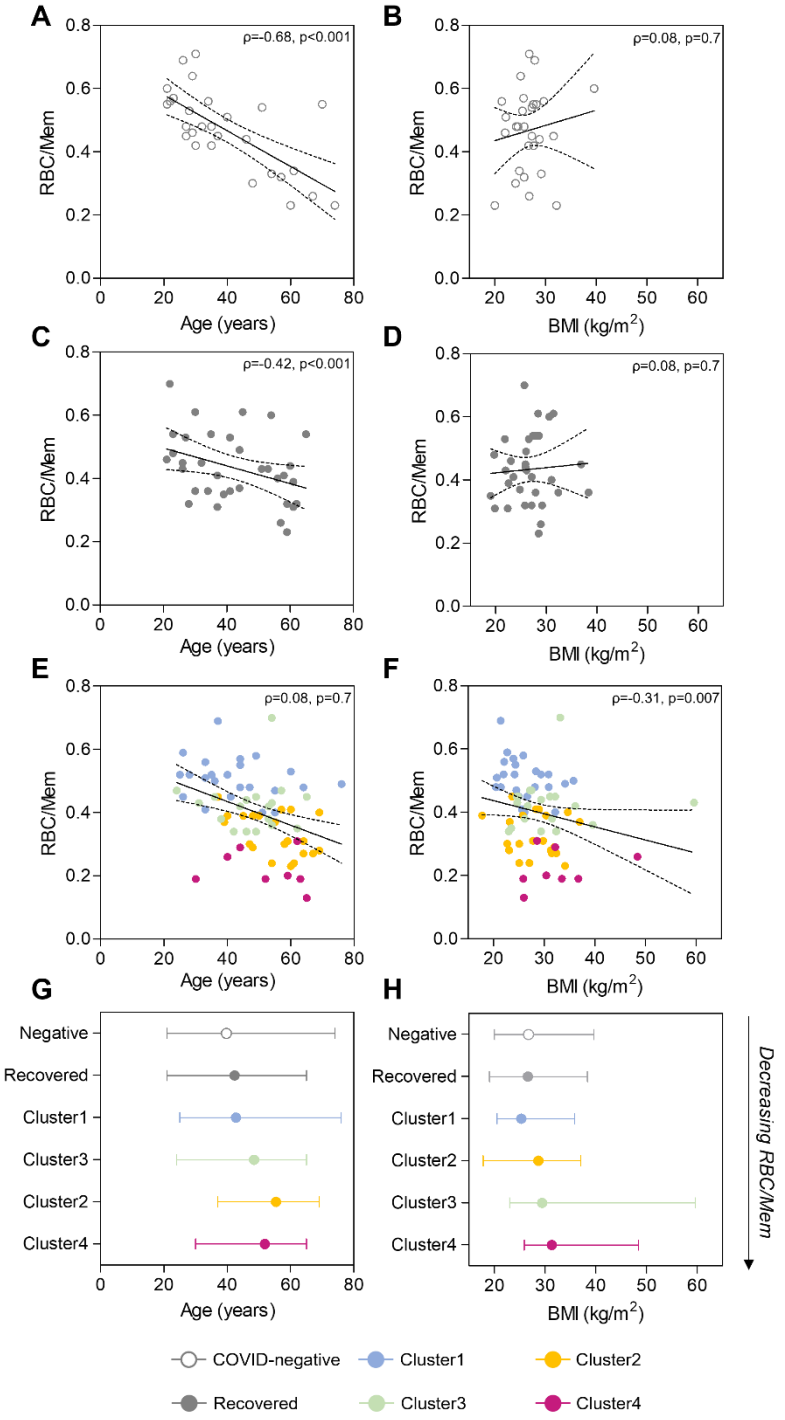


Figure E3. Age and BMI Relationships with MRI RBC/Mem.
^{129}Xe MRI red blood cell to interstitial membrane (RBC/Mem) ratio was significantly correlated with age for (A) COVID-negative, (C) fully recovered, and (E) long COVID groups. RBC/Mem was not significantly correlated with body mass index (BMI) in COVID-negative (B) nor recovered (D) groups. RBC/Mem and BMI were significantly but weakly correlated for all long COVID participants (F). Long COVID clusters shown with coloured points. (D, H) Points show mean age or BMI with range, for groups arranged from top to bottom with decreasing RBC/Mem, demonstrating overlapping age and BMI ranges across all groups.



REFERENCES

1. He M, Robertson SH, Kaushik SS, et al. Dose and pulse sequence considerations for hyperpolarized (129)Xe ventilation MRI. *Magn Reson Imaging* 2015: 33(7): 877-885.

2. Niedbalski PJ, Hall CS, Castro M, et al. Protocols for multi-site trials using hyperpolarized (129) Xe MRI for imaging of ventilation, alveolar-airspace size, and gas exchange: A position paper from the (129) Xe MRI clinical trials consortium. *Magn Reson Med* 2021: 86(6): 2966-2986.

3. Kaushik SS, Robertson SH, Freeman MS, et al. Single-breath clinical imaging of hyperpolarized (129)Xe in the airspaces, barrier, and red blood cells using an interleaved 3D radial 1-point Dixon acquisition. *Magn Reson Med* 2016: 75(4): 1434-1443.

4. Niedbalski PJ, Willmering MM, Thomen RP, et al. A single-breath-hold protocol for hyperpolarized (129) Xe ventilation and gas exchange imaging. *NMR Biomed* 2023: e4923.

5. Niedbalski PJ, Lu J, Hall CS, et al. Utilizing flip angle/TR equivalence to reduce breath hold duration in hyperpolarized (129) Xe 1-point Dixon gas exchange imaging. *Magn Reson Med* 2022: 87(3): 1490-1499.

6. Ruppert K, Amzajerdian F, Hamedani H, et al. Assessment of flip angle-TR equivalence for standardized dissolved-phase imaging of the lung with hyperpolarized 129Xe MRI. *Magn Reson Med* 2019: 81(3): 1784-1794.

7. Leewiwatwong S, Lu J, Dummer I, et al. Combining neural networks and image synthesis to enable automatic thoracic cavity segmentation of hyperpolarized (129)Xe MRI without proton scans. *Magn Reson Imaging* 2023.

8. Wang Z, Robertson SH, Wang J, et al. Quantitative analysis of hyperpolarized (129) Xe gas transfer MRI. *Med Phys* 2017: 44(6): 2415-2428.

9. Davies DL, Bouldin DW. A cluster separation measure. *IEEE Trans Pattern Anal Mach Intell* 1979(2): 224-227.

10. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comp Appl Math* 1987: 20: 53-65.

11. Dunn JC. Well-separated clusters and optimal fuzzy partitions. *Journal of Cybernetics* 1974: 4(1): 95-104.

12. Plummer JW, Willmering MM, Cleveland ZI, et al. Childhood to adulthood: Accounting for age dependence in healthy-reference distributions in (129) Xe gas-exchange MRI. *Magn Reson Med* 2023: 89(3): 1117-1133.

13. Mummy D, Swaminathan A, Bier E, et al. Hyperpolarized 129Xe MRI and spectroscopy in healthy control subjects reveals age-related changes in measurements of pulmonary gas exchange. *Proc Intl Soc Magn Reson Med* 2022: 30(1168).