Article

Pulmonary and renal long COVID at two-year revisit

Long COVID at One- and Two-year Revisits

Follow up schedule → Day 30 144 patients → Year 1(R1) 73 patients → Year 2(R2) 57 patients

Organ long COVID - based on clinical data

Renal Long COVID: eGFR<90mL/min/1.73m²

Pulmonary Long COVID: Fibrous stripe lesions

Hepatic long COVID Abnormal liver function indicators

	Renal	Pulmonary	Hepatic
Rate at R1:	15.2%	8.7%	32.6%
Rate at R2:	23.9%	17.8%	19.6%

Machine learning - based on data before Day 30

×144 → Before Day 30 → ×81 Proteomics & Metabolomics → ×66 Model group / ×15 Validation group → Machine learning for Pulmonary and Renal Long COVID at R1 → Evaluation → Clinical indicator model / Multi-omics model / Clinical indicator and multi-omics model

Pulmonary and Renal Long COVID Related molecules

Renal Long COVID — Proenkephalin (PENK) ↑ / MASP1 ↓

Pulmonary Long COVID — ATR (ataxia telangiectasia and Rad3-related protein) ↑ / SFTPB ↓

Jing Wang, Xiao Liang, Yufen Zheng, ..., Jiaqin Xu, Tiannan Guo, Bo Shen

chenhx@enzemed.com (H.C.)
lij@enzemed.com (J.L.)
xujq@enzemed.com (J.X.)
guotiannan@westlake.edu.cn (T.G.)
shenb@enzemed.com (B.S.)

Highlights

A two-year molecular landscape resource for COVID patients

A multi-omics model to predict pulmonary and renal long COVID one year ahead

Host response analysis for pulmonary and renal long COVID

Article

Pulmonary and renal long COVID at two-year revisit

Jing Wang,[1,7,8,9] Xiao Liang,[2,3,4,9] Yufen Zheng,[1,7,8,9] Yi Zhu,[2,3,4,9] Kai Zhou,[1,9] Xiaomai Wu,[1,9] Rui Sun,[2,3,4] Yifan Hu,[5] Xiaoli Zhu,[1] Hongbo Chi,[1] Shanjun Chen,[5] Mengge Lyu,[2,3,4] Yuting Xie,[2,3,4] Xiao Yi,[5] Wei Liu,[5] Xue Cai,[2,3,4] Sainan Li,[2,3,4] Qiushi Zhang,[5] Chunlong Wu,[5] Yingqiu Shi,[2,3,4] Donglian Wang,[1] Minfei Peng,[1] Ying Zhang,[1] Huafen Liu,[6] Chao Zhang,[6] Sheng Quan,[6] Ziqing Kong,[6] Zhouyang Kang,[6] Guangjun Zhu,[1] Hongguo Zhu,[1] Shiyong Chen,[1] Junbo Liang,[1] Hai Yang,[1] Jianxin Pang,[1] Yicheng Fang,[1] Haixiao Chen,[1,*] Jun Li,[1,7,8,*] Jiaqin Xu,[1,7,8,*] Tiannan Guo,[2,3,4,*] and Bo Shen[1,7,10,10,*]

SUMMARY

This study investigated host responses to long COVID by following up with 89 of the original 144 cohorts for 1-year ($N = 73$) and 2-year visits ($N = 57$). Pulmonary long COVID, characterized by fibrous stripes, was observed in 8.7% and 17.8% of patients at the 1-year and 2-year revisits, respectively, while renal long COVID was present in 15.2% and 23.9% of patients, respectively. Pulmonary and renal long COVID at 1-year revisit was predicted using a machine learning model based on clinical and multi-omics data collected during the first month of the disease with an accuracy of 87.5%. Proteomics revealed that lung fibrous stripes were associated with consistent down-regulation of surfactant-associated protein B in the sera, while renal long COVID could be linked to the inhibition of urinary protein expression. This study provides a longitudinal view of the clinical and molecular landscape of COVID-19 and presents a predictive model for pulmonary and renal long COVID.

INTRODUCTION

As of March 19, 2022, more than 460 million people worldwide had been affected by COVID-19, with more than 98.5% of these cases cured.[1] Approximately 70% of recovered patients experienced long COVID symptoms 6 months after the disease,[2] becoming a global concern.

Long COVID affects nearly all organs,[3,4] with the most reported symptoms including chest pain, dyspnea, fatigue, smell or taste disorders, and neurocognitive effects such as brain fog, memory difficulty, anxiety or confusion, and cognitive impairment.[5,6] These symptoms emerged during or after the COVID-19 pandemic and persisted for several months.[7]

One study reported that long-lasting lung lesions were present up to 12 months after discharge,[8] with 24% of the enrolled patients ($N = 20$) exhibiting unresolved radiological changes, mainly in the form of ground-glass opacity (GGO) lesions. Renal long COVID was reported up to 6 months after discharge,[9] with 35% of the cases ($N = 487$) having abnormal estimated glomerular filtration rate (eGFR) and 22% exhibiting persistent renal dysfunctions since the disease onset. Nevertheless, the longer-term consequences of these long COVID symptoms are unclear, owing to the limited monitoring time span. In addition, the molecular mechanisms underlying pulmonary and renal long COVID remain elusive. Therefore, early prevention of pulmonary and renal long COVID remains challenging.

In this study, we leveraged the results from a cohort of 144 patients with COVID-19 previously enrolled during the first pandemic in China[10–12] to develop a predictive model for pulmonary and renal long COVID. These data encompassed cutting-edge mass spectrometry (MS)-based multi-omics technologies (proteomics and metabolomics), clinical parameters, and chest computed tomography (CT) scan images. We enrolled 89 of the 144 original patients who consented to 1-year and/or 2-year follow-up visits and integrated their multi-omics data with clinical manifestations. These data were used in developing machine learning models to predict pulmonary and renal long COVID that emerged 1 year after disease onset, highlighting potential regulatory factors associated with them.

[1]Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, China
[2]Center for Intelligent Proteomics, Westlake Laboratory of Life Sciences and Biomedicine, Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang Province, China
[3]Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang Province, China
[4]Research Center for Industries of the Future, Westlake University, Hangzhou, Zhejiang, China
[5]Westlake Omics (Hangzhou) Biotechnology Co., Ltd, Hangzhou 310024, China
[6]Calibra Lab at DIAN Diagnostics, 329 Jinpeng Street, Hangzhou 310030, Zhejiang Province, China
[7]Key Laboratory of System Medicine and Precision Diagnosis and Treatment of Taizhou, Taizhou, Zhejiang, China
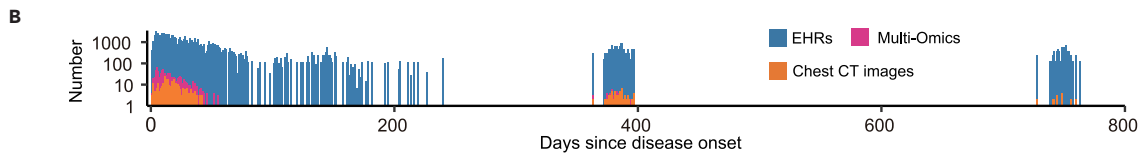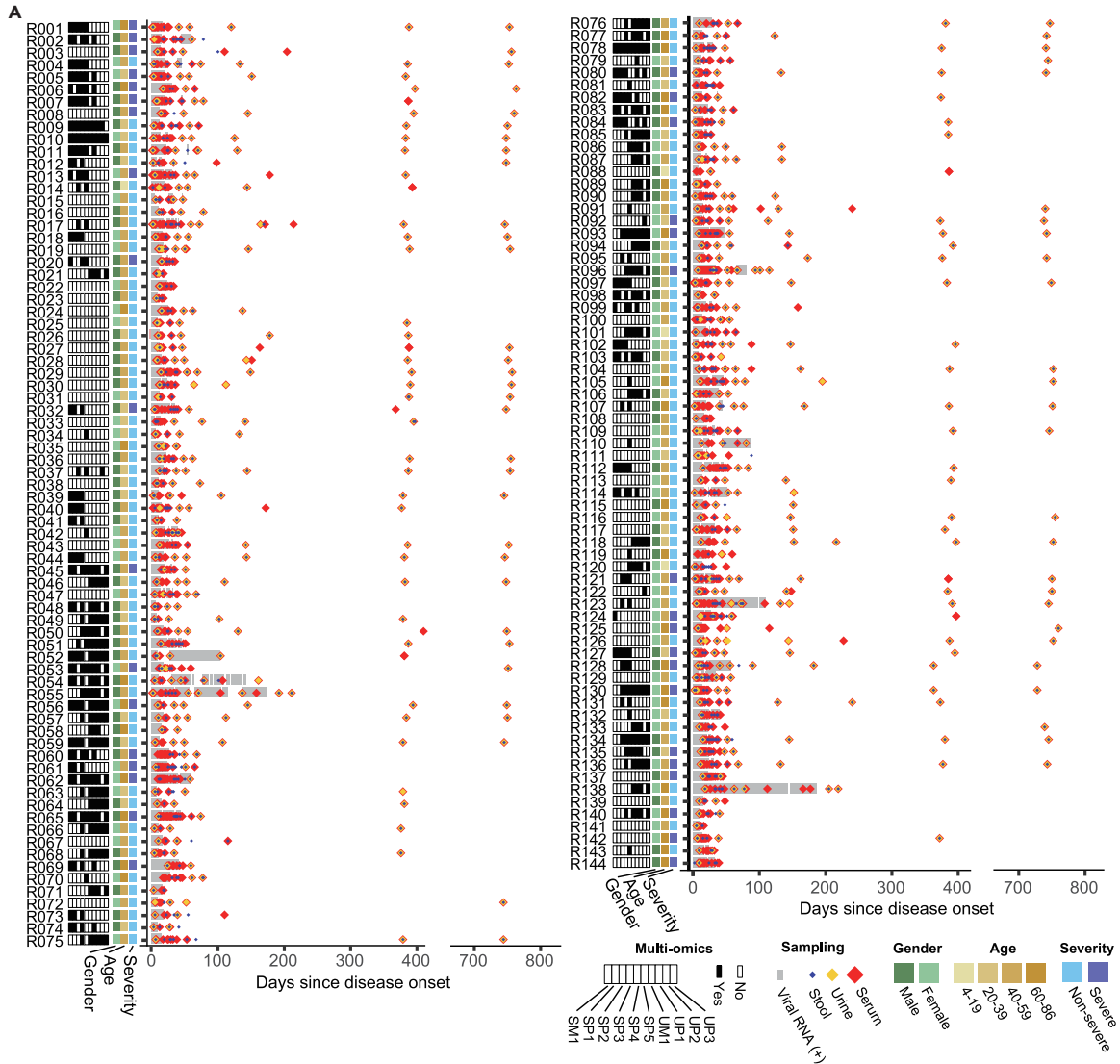[8]Taizhou Institute of Medicine, Health and New Drug Clinical Research, Taizhou, Zhejiang, China
[9]These authors contributed equally
[10]Lead contact
*Correspondence: chenhx@enzemed.com (H.C.), lij@enzemed.com (J.L.), xujq@enzemed.com (J.X.), guotiannan@westlake.edu.cn (T.G.), shenb@enzemed.com (B.S.)
https://doi.org/10.1016/j.isci.2024.110344

A

B

EHR	Viral RNA	Kidney	Liver	Intestine	Infection	Lipid	Glucose	Immunity	Electrolyte	Coagulation	Heart	Complete blood count	Lung
Types	4	59	15	7	6	7	3	25	10	6	8	22	10
Datapoints	2144	19074	13660	3506	2926	5520	1675	15218	8509	2508	1892	23298	9040

Chest CT

Tests 435

MS-based Multi-omics

Dataset	SM1	SP1	SP2	SP3	SP4	SP5	UM1	UP1	UP2	UP3
Patients	46	41	37	66	54	50	64	50	28	46
Samples	46	41	108	231	299	50	64	50	56	46
IDs	941	894	788	22	337	1384	1033	3737	4161	31

■ Serum metabolites
■ Serum proteins
■ Urine metabolites
■ Urine proteins

Figure 1. Experimental outline and patient cohort

(A) Baseline patient information and molecular samplings. The x axis shows the assessment times since the disease onset, and the y axis displays the 144 patients from the original study cohort. A patient was enrolled if their SARS-CoV-2 throat/sputum nucleic acid test was positive on day 1. The positive and negative results of the nucleic acid tests are shown with red and gray bars, respectively. Three types of samples (blood, urine, and stool) were collected, and twelve MS-based multi-omics were performed, as indicated in the study outline.

(B) Summary of molecular samplings. The upper panel shows the time-resolved number of three dataset types (EHRs, omics, and chest CT images) across the 763 days of the study. The lower panel shows the subtypes of each dataset and the sampling number for each subtype. EHRs, electronic hospital records; CT, computed tomography.

RESULTS

An integrated and temporal COVID-19 molecular landscape resource

During the first COVID-19 pandemic, 144 patients with COVID-19 from the Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University (China) were enrolled (Figure 1A; Table S1).[10–12] The patients were admitted on day 4–10 (median, 6) from disease onset. They were clinically examined throughout the disease course and discharged on day 20–36 (median, 26). In this study, we recruited 73 and 58 patients for 1-year and 2-year follow-up visits, respectively (Table S2; Figure S1). We used 182 of the previous clinical indices that were recorded during the index hospitalization and prior clinical lab results of blood, urine, and stool, resulting in 13 sets of clinical assessments of physiological conditions (Figure 1B).[10–12] We used chest CT image reports interpreted by radiologists. At the 1-year and 2-year follow-up visits, the patients underwent 11 sets of clinical assessments of physiological conditions (except for RNA and pulmonary function-related laboratory indices), chest CT, and pulmonary function. At the 1-year follow-up visit, serum and urine samples from 37 to 28 patients, respectively, were used for MS-based proteomic analyses. However, at the 2-year follow-up, samples did not undergo multi-omics analysis (Figure S4; Table S3).[10–12] We incorporated data from seven prior multi-omics datasets[10–12]; creating two additional serum proteomics datasets and one urine proteomics dataset for analysis. Altogether, we analyzed multi-omics data on 89 patients with follow-up data, including 174 serum proteomes from 63 patients, 90 urine proteomes from 37 patients, 28 serum metabolomes from 28 patients, and 40 urine metabolomes from 40 patients. The inclusion criteria for each cohort are shown in Figure S1 and Table S4.

Clinical characteristics of long COVID

Recovered COVID-19 patients showed differences in symptoms at the 1-year (R1) and 2-year (R2) revisits. R2 patients had a higher incidence of sleep disorders (34.8% vs. 19.6%, $p = 0.101$), palpitations (26.1% vs. 10.9%, $p = 0.060$), alopecia (28.3% vs. 10.9%, $p = 0.036$), and erythra (21.7% vs. 8.7%, $p = 0.082$) compared to R1 patients (Table 1).

Of the 46 patients available for both R1 and R2 visits, 30.4% and 32.6% exhibited abnormal pulmonary function and chest CT findings at R1, respectively. These percentages decreased to 23.9% and 24.4% at R2, respectively ($p = 0.482$; $p = 0.529$). Notably, fibrous stripes were observed more in R2 patients than to R1 patients (17.8% vs. 8.7%, $p = 0.200$). Interestingly, abnormal renal function was observed in 15.2% of patients at R1 but in 23.9% at R2 ($p = 0.293$), shown by the significantly decreased eGFR ($p = 0.010$) at R2. This indicates potential cumulative kidney damage in the revisited patients. In contrast, only 19.6% of the patients exhibited liver abnormalities at R2, an improvement from R1 (32.6%) ($p = 0.235$), suggesting the recovery of hepatic long COVID (Table 1). Consequently, we did not focus specifically on liver biomarkers in this study. Our data showed that patients who were female, over 50 years old, with severe symptoms or underlying diseases, or a high BMI (>25 kg/m²) had higher rates of 1- and 2-year pulmonary and renal long COVID symptoms (Figure S3). Notably, chest CT findings were common in severe cases, with more than 70% and 60% of the observed cases at R1 and R2, respectively. Their ratios of chest CT findings in the severe patients were 1-fold and two-times higher than the non-severe ones at R1 and R2, respectively.

While long COVID is characterized by a panoply of symptoms and signs, data analysis from our longitudinal cohort showed that the most frequently affected organs were lungs and kidneys in the long term, showing objective and measurable features; therefore, we focused on pulmonary and renal long COVID in this study. Nevertheless, we could not exclude the possibility that other uncharacterized factors might play a role in the subsequent lung and kidney abnormalities observed, such as infections (e.g., influenza and other viruses) or non-COVID-related illnesses acquired during the study period. We also compared the proportion of other long COVID symptoms between patients with and without pulmonary or renal long COVID among the 73 patients who participated in the R1 revisit. We found that patients with pulmonary long COVID had a higher proportion of sleep disorders, alopecia, and joint pain, while patients with renal long COVID patients had a higher proportion of fatigue, palpation, joint pain, and dizziness. However, the differences between the groups in all long COVID symptoms were not statistically significant (Table S12).

Multi-omics and clinical data were used to create a prediction model for pulmonary and renal long COVID

To predict the incidence of pulmonary and renal long COVID, we developed a machine learning-based model using clinical and multi-omics data collected from the original cohort during the index hospitalization and within 30 days of initial illness.

Among the 73 patients who were followed-up at R1, we excluded seven for the following reasons: two patients with a self-reported medical history of chronic renal diseases, two patients under 18 years old, two patients without available pulmonary function test reports, and one patient whose chest CT was performed 1 month before R1. Finally, 66 patients who underwent chest CT and kidney function assessments at R1 were included in this study (Figure 2A). Patients with chest CT abnormalities or/and renal dysfunctions at R1 were classified as the pulmonary or/and renal long COVID (LC) group ($N = 26$), while patients without chest CT abnormalities and renal dysfunctions were defined as

Table 1. Summary of one- and two-year long COVID

	Overall			Overlapped		
	R1 (*N* = 73)	R2 (*N* = 57)	*p* value	R1 (*N* = 46)	R2 (*N* = 46)	*p* value
Symptom						
Fatigue, n (%)	28 (38.4)	21 (36.8)	0.860	15 (32.6)	19 (41.3)	0.388
Sleep disorder, n (%)	18 (24.7)	22 (38.6)	0.088	9 (19.6)	16 (34.8)	0.101
Dysosmia, n (%)	0 (0.0)	2 (3.5)	0.190[a]	0 (0.0)	2 (4.3)	0.475[b]
Chest pain, n (%)	11 (15.1)	10 (17.5)	0.704	4 (8.7)	7 (15.2)	0.335
Palpitation, n (%)	12 (16.4)	15 (26.3)	0.168	5 (10.9)	12 (26.1)	0.060
Alopecia, n (%)	10 (13.7)	17 (29.8)	0.025	5 (10.9)	13 (28.3)	0.036
Erythra, n (%)	8 (11.0)	11 (19.3)	0.182	4 (8.7)	10 (21.7)	0.082
Myalgia, n (%)	11 (15.1)	3 (5.3)	0.074	4 (8.7)	2 (4.3)	0.677[a]
Joint pain, n (%)	19 (26.0)	14 (24.6)	0.849	11 (23.9)	10 (21.7)	0.804
Decreased appetite, n (%)	2 (2.7)	5 (8.8)	0.263[b]	1 (2.2)	3 (6.5)	0.609[b]
Taste disorder, n (%)	3 (4.1)	0 (0.0)	0.337[b]	1 (2.2)	0 (0.0)	1.000[a]
Dysphagia, n (%)	15 (20.5)	7 (12.3)	0.212	8 (17.4)	7 (15.2)	0.778
Dizzy, n (%)	18 (24.7)	10 (17.5)	0.328	12 (26.1)	6 (13.0)	0.115
Headache, n (%)	11 (15.1)	6 (10.5)	0.446	5 (10.9)	4 (8.7)	1.000[b]
Diarrhea and vomiting, n (%)	7 (9.6)	5 (8.8)	0.873	3 (6.5)	4 (8.7)	1.000[b]
Low fever, n (%)	5 (6.8)	3 (5.3)	0.995[b]	4 (8.7)	2 (4.3)	0.673[b]
Chest CT radiography						
Abnormality, n (%)	25 (34.2)	16 (28.6)	0.620	15 (32.6)	11 (24.4)	0.529
CT score, median (IQR)	0.0 (0.0–2.5)	0.0 (0.0–2.8)	0.982[c]	0.0 (0.0–2.0)	0.0 (0.0–1.0)	0.657[c]
Involved lobes, median (IQR)	0.0 (0.0–1.0)	0.0 (0.0–1.8)	0.931[c]	0.0 (0.0–1.0)	0.0 (0.0–0.5)	0.590[c]
PO/GGO, n (%)	18 (24.7)	14 (25.0)	0.964	13 (28.3)	10 (22.2)	0.508
Consolidation, n (%)	0 (0.0)	0 (0.0)	NA	0 (0.0)	0 (0.0)	NA
Fibrous stripe, n (%)	10 (13.7)	13 (23.2)	0.162	4 (8.7)	8 (17.8)	0.200
Organ function						
Pulmonary function						
Abnormality, n (%)	21 (29.2)	13 (22.8)	0.416	14 (30.4)	11 (23.9)	0.482
Impaired obstructive ventilation, n (%)	11 (15.3)	4 (7.0)	0.146	5 (10.9)	4 (8.7)	1.000[b]
FEV1, %pred, median (IQR)	93.0 (84.0–103.2)	95.0 (88.0–101.0)	0.435[c]	93.0 (84.8–100.5)	95.5 (89.2–100.5)	0.362[c]
FEV1/FVC, %pred, median (IQR)	99.0 (94.0–104.0)	100.0 (96.0–104.0)	0.678[c]	99.0 (94.2–103.8)	100.0 (96.0–104.8)	0.455[c]
Impaired restrictive ventilation, n (%)	5 (6.9)	4 (7.0)	1.000[b]	4 (8.7)	3 (6.5)	1.000[b]
FVC, %pred, median (IQR)	93.5 (86.5–105.0)	94.0 (88.0–105.0)	0.807[c]	92.0 (85.5–105.0)	93.5 (88.0–105.0)	0.752[c]
TLC, %pred, median (IQR)	100.0 (91.5–107.0)	97.0 (87.5–107.5)	0.585[c]	98.5 (91.2–108.8)	97.5 (89.8–108.0)	0.656[c]
RV, %pred, median (IQR)	109.0 (98.0–124.0)	110.0 (100.0–119.0)	0.833[c]	106.5 (98.0–118.8)	110.0 (100.2–118.8)	0.524[c]
RV/TLC, %pred, median (IQR)	112.0 (102.5–122.0)	115.0 (104.0–122.5)	0.697[c]	109.0 (102.2–119.5)	111.0 (102.0–121.2)	0.619[c]
Diffusion dysfunction, n (%)	12 (16.7)	10 (17.5)	0.895	10 (21.7)	8 (17.4)	0.599
DLCO, %pred, median (IQR)	95.5 (88.2–107.0)	97.0 (87.0–110.0)	0.625[c]	97.5 (83.0–109.5)	96.0 (84.2–108.0)	0.897[c]
KCO, %pred, median (IQR)	100.5 (90.0–112.2)	104.0 (93.0–115.0)	0.458[c]	100.5 (93.8–113.0)	100.5 (93.0–112.8)	0.873[c]
Small airway dysfunction, n (%)	16 (22.2)	5 (8.8)	0.040	9 (19.6)	5 (10.9)	0.246
FEF75, %pred, median (IQR)	76.5 (56.0–101.0)	77.0 (59.0–100.0)	0.630[c]	77.5 (56.2–94.8)	82.5 (58.2–104.2)	0.494[c]
FEF50, %pred, median (IQR)	88.0 (68.8–110.0)	94.0 (77.0–109.0)	0.372[c]	90.0 (69.5–108.2)	94.5 (78.2–109.0)	0.292[c]
FEF25-75, %pred, median (IQR)	89.0 (68.8–107.2)	95.0 (75.0–110.0)	0.313[c]	90.5 (69.2–103.8)	96.0 (75.8–114.5)	0.259[c]

Table 1. Continued

	Overall			Overlapped		
	R1 (N = 73)	R2 (N = 57)	p value	R1 (N = 46)	R2 (N = 46)	p value
Renal function						
Abnormality, n (%)	10 (13.7)	12 (21.1)	0.267	7 (15.2)	11 (23.9)	0.293
eGFR, mL/min/1.73m^2, median (IQR)	112.0 (99.0–120.0)	98.0 (92.0–110.0)	<0.001[c]	109.0 (98.0–117.0)	98.0 (90.0–110.0)	0.010[c]
Liver function						
Abnormality, n (%)	22 (30.1)	11 (19.3)	0.228	15 (32.6)	9 (19.6)	0.235
ALT, U/L, median (IQR)	17.0 (13.0–26.0)	18.0 (14.0–29.0)	0.260[c]	16.5 (13.2–25.8)	18.0 (14.0–28.2)	0.434[c]
AST, U/L, median (IQR)	20.0 (18.0–24.0)	20.0 (18.0–26.0)	0.951[c]	22.0 (18.0–26.2)	21.0 (18.0–26.8)	0.781[c]
ALP, U/L, median (IQR)	72.0 (57.0–87.0)	72.0 (59.0–87.0)	0.844[c]	73.5 (61.2–87.8)	71.0 (57.5–80.5)	0.216[c]
GGT, U/L, median (IQR)	20.0 (15.0–27.0)	21.0 (15.0–32.0)	0.645[c]	20.0 (16.0–26.8)	20.0 (15.0–28.5)	0.873[c]
Total bilirubin, μmol/L, median (IQR)	14.6 (12.2–18.6)	11.5 (9.6–15.1)	<0.001[c]	14.9 (12.2–20.4)	11.4 (9.6–15.7)	0.004[c]

R1, one-year revisit; R2, two-year revisit; FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; TLC, total lung capacity; RV, residual volume; DLCO, diffusing capacity of lung carbon monoxide; KCO, carbon monoxide transfer coefficient; FEF, forced expiratory flow; eGFR, estimated glomerular filtration rate; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; GGT, gamma-glutamine transferase.
The percentages in Table 1 were calculated based on the patient number at different timepoints. Of the 73 patients at R1, one did not participate in pulmonary function examination at R1. Another patient contains missing values in its pulmonary function tests but we were still able to evaluate the related function from the remaining data. Of the 57 patients at R2, one patient did not participate in chest CT examination. Other p values were obtained by Pearson χ^2 test.
[a]Fisher exact test.
[b]Continuity correction Pearson χ^2 test.
[c]Mann-Whitney U test.

the recovered (RC) group (N = 40). We used 75% of the cohort (50/66) as the training set (randomly selected) and 25% (16/66) as the internal test set (Figure S2A). For the clinical model, patients were randomly divided into clinical training (N = 34) and validation (N = 16) sets (Figure S2B). The XGBoost modeling identified a prognostic panel comprising nine clinical indicators (Figure 2E and STAR methods), with an area under the curve (AUC) of 0.783 (95% confidence interval 0.555–1.000) and an accuracy (ACC) of 0.812 (Figure 2B). Patients with omics data were grouped into 10 multi-omics training sets for the omics model (Figure S3A). We included no more than five features from each model (STAR methods), ranked by XGBoost importance (Figure S2D; Table S11) and assembled them into a multi-omics model comprising 36 molecules, including two serum metabolites (Figure 2F), five urine metabolites (Figure 2F), 20 serum proteins (Figure 2G), and nine urine proteins (Figure 2H). Applying the multi-omics model to the internal test set resulted in an AUC of 0.900 (95% confidence interval, 0.745–1.000) and an ACC of 0.812 (Figure 2C), outperforming the clinical model. Integrating multi-omics with clinical indicators increased the AUC to 0.983 (95% confidence interval, 0.921–1.000) and the ACC to 0.875 (Figures 2D and S2C). Two patients were identified incorrectly (Figure 2I). Patient R046 was identified as having long COVID due to the presence of PO (patchy opacity)/GGO lesions on the chest CT and an abnormal eGFR at R1 (Figure S2E). Moreover, we observed that his fasting blood glucose level on admission was 7.76 mmol/L, suggesting a potential pre-existing diabetes condition before the SARS-CoV-2 infection; however, this was not rigorously confirmed medically. The complexity of the patient's medical history might have distorted our model. Patient R097 had elevated serum anti-streptolysin O (ASO) levels (Figure S2F) and lung restrictive ventilation dysfunction at the revisit, suggesting that a streptococcal infection post-COVID-19 might have contributed to the inaccurate prediction.[13]

Validation of key serum molecules and clinical indices for the model

To validate the model, an independent cohort of 15 patients was used, consisting of 7 patients who participated in the 1-year revisit but were not included in the modeling group and 8 patients who did not participate in the 1-year revisit but had lung and/or kidney long COVID outcomes without self-reported symptoms at R1. They were enrolled as an external test set (Figure S2A). Their data on prioritized clinical indicators and serum proteins during the 30 days following disease onset were collected to evaluate the model's performance (Table S6 and STAR methods), and their long COVID status was assessed from revisits on days 386–487 (median, 397) after COVID-19. The clinical indicators exhibited an AUC of 0.685 and an ACC of 0.667 (Figures 3A and 3B). Simultaneously, the targeted proteomic dataset of 20 serum protein candidates generated an AUC of 0.800 and an ACC of 0.750 (Figures 3A and 3C), suggesting that the prioritized serum proteins may predict pulmonary and renal long COVID. Three male patients (R032, R055, and R058) were incorrectly predicted using serum proteins (Figure 3C). Patient R032 had a history of hepatitis B infection and an underlying thyroid disease. Patient R058 had a history of hypertension and had received long-term antihypertensive medication, while patient R055 was asymptomatic.

Five biological functions were significantly enriched in the serum protein panels (Figure 3D), including acute inflammation and fibrinolysis, which are frequently documented in COVID-19.[14] Our data suggest that these changes also contributed to long-term injuries.
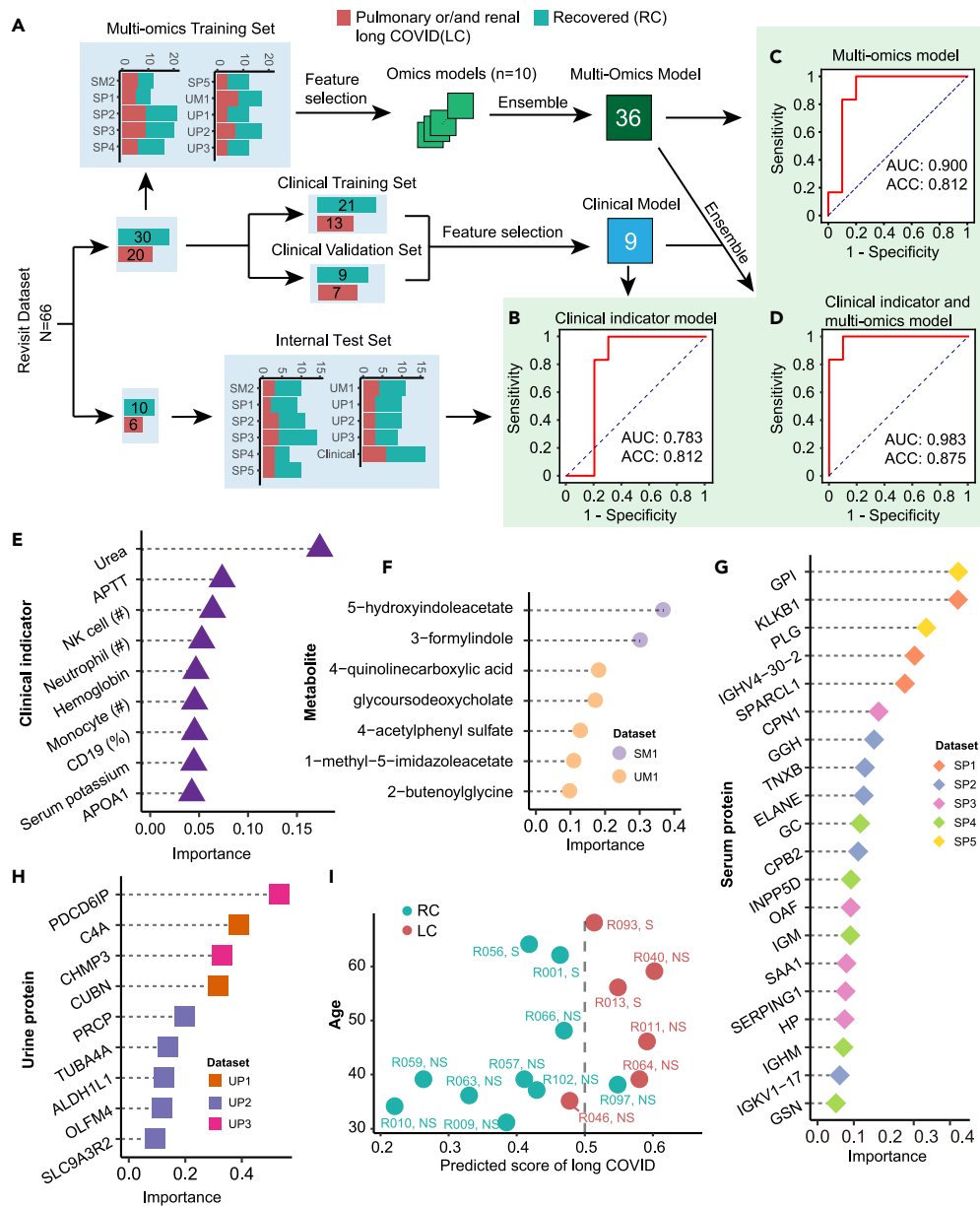
Figure 2. Machine learning-based prediction of long COVID

(A) Overview of patient classification, machine learning, and performance modules of the prediction model. The x axis of the bar plots denotes the number of patients analyzed per multi-omics panel. The prediction models were created based on the multi-omics and clinical data within 30 days after the disease onset, and 1-year pulmonary and renal long COVID outcome.

(B) Receiver operating characteristic (ROC) curve for predicting long COVID using the clinical indicator model of the internal test set.

(C) ROC curve for predicting long COVID using the multi-omics model of the internal test set.

(D) ROC curve for predicting long COVID using the clinical indicator and multi-omics integrated model of the internal test set.

(E) Nine clinical indicators prioritized by the machine learning model.

(F) The machine learning model prioritized seven serum and urine metabolites.

(G) Twenty serum proteins prioritized by the machine learning model.

(H) Nine urine proteins prioritized by the machine learning model.

(I) Performance of the integrated model on the test set of 16 patients. S, severe; NS, non-severe. LC, patients with pulmonary or/and renal long COVID at 1-year revisit; RC, patients without pulmonary and renal long COVID at 1-year revisit. The predicted score of long COVID ranges from 0 to 1, indicating the probability of development of long COVID. A score higher than 0.5 suggests the patient is more likely to have long COVID. Importance is a designated parameter in the XGBoost model to assess the contribution of each feature.
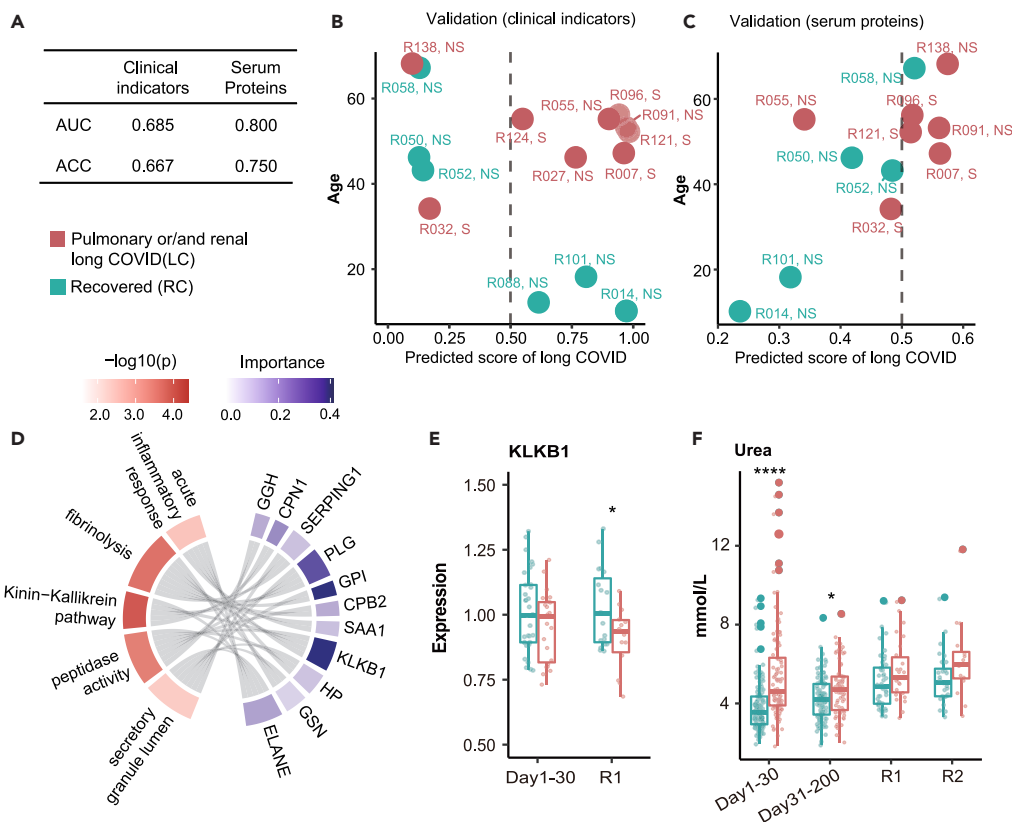
Figure 3. Validation of long COVID prediction

(A) AUC values of long COVID prediction using serum proteins and clinical indicators in the external test cohort.

(B and C) Performance of clinical indicators on the validation set of 15 patients and serum proteins on the validation set of 12 patients. S, severe; NS, non-severe.

(D) Key functions enriched from the serum proteins. Annotation bars of the proteins and the enriched functions denote the feature importance and significance, respectively.

(E and F) Temporal expression of KLKB1 and blood urea at different stages. Data are represented as median (interquartile range). The signs of significance were described in p values, *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.005$; ****, $p < 0.001$. LC, patients with pulmonary or/and renal long COVID at 1-year revisit; RC, patients without pulmonary and renal long COVID at 1-year revisit.

Coagulation-related markers, such as plasminogen (PLG) and carboxypeptidase B2 (CPB2), may also indicate liver functions. Particularly, plasma kallikrein (KLKB1) expression in the long COVID (LC) group was comparable to that in the recovered (RC) group during the disease stage but significantly lower than that in the RC group at R1 (Figure 3E). Serum KLKB1 levels are a contributing factor to lung injuries.[15] Our data highlight it as a promising biomarker for pulmonary long COVID. For clinical indicators, blood urea, prioritized as the marker with the best predictive power (Figure 3F), was significantly overexpressed in patients with LC and remained high during the monitoring stage. Urea is an established marker of renal function associated with pulmonary injuries in COVID-19.[16] Our data indicate that sustained high blood urea levels may be related to both pulmonary and renal long COVID.

Persistent fibrous stripes in pulmonary long COVID

Next, we evaluated lung injuries based on the chest CT images. Patchy opacity (PO) and ground-glass opacity (GGO) were the most prevalent lesions observed during hospitalization (Figure 4A). The number of patients with PO/GGO and chest consolidation decreased from T1 (on admission) to T3 (before discharge), while the number of patients with fibrous stripes increased. These observations align with those of other studies, which report that chest lesions, especially fibrotic lesions, are not adequately resolved by discharge.[12] At R1, all types of lesions showed significant remission, with no consolidation observed. Based on the overlapping patients, the ratios of PO/GGO lesions decreased from 28.3% at R1 to 22.2% at R2 ($p = 0.508$), while the percentages of fibrous stripes increased from 8.7% at R1 to 17.8% at R2 ($p = 0.200$). Notably, the fibrous stripes persisted at R2 in all cases, as shown in Table S2. Surprisingly, three patients without fibrous stripes from T1 to R1 developed fibrous stripes at R2. A detailed analysis of temporal chest CT images is elaborated in Figure 5A. Based on the radiological manifestations, the fibrous stripes were identified as persistent pulmonary lesions. Among the pulmonary functions, only small airway dysfunction exhibited a significantly lower incidence at R2 than at R1 ($p = 0.04$; Table 1). There were 63 patients without fibrous stripes and 10 patients with fibrous stripes at R1. We classified them into the NFIB and FIB groups, respectively. Examples of their chest CT images obtained at R1 are
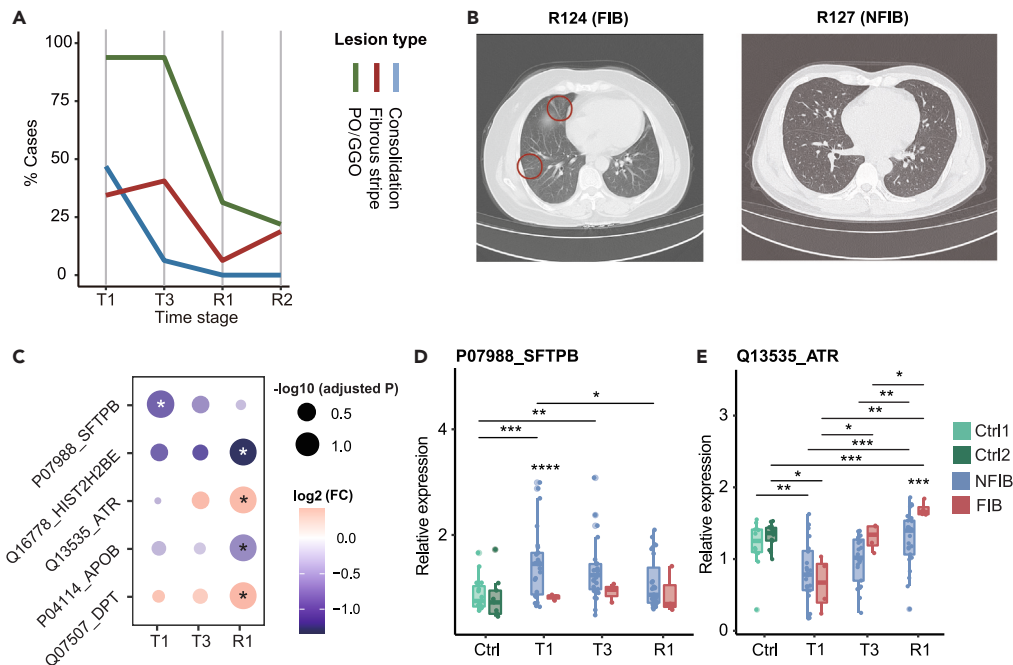
Figure 4. Pulmonary long COVID in the recovered patients

(A) Development of chest CT lesions from T1 to R2.

(B) Examples of chest CT images for the NFIB and FIB patients.

(C) Fold changes of serum DEPs between the NFIB and FIB group at T1, T3, and R1. The signs of significance were described in adjusted *p* values.

(D and E) Temporal serum SFTPB and ATR expression at T1, T3, and R1. Data are represented as median (interquartile range). The signs of significance were described in *p* values, *, *p* < 0.05; **, *p* < 0.01; ***, *p* < 0.005; ****, *p* < 0.001. DEP, differentially expressed protein; NFIB, revisited patients at R1 with no fibrous stripes; FIB, revisited patients at R1 with fibrous stripes; Ctrl1, non-COVID-19 individuals with no fibrous stripes; Ctrl2, non-COVID-19 individuals with fibrous stripes. *p* values were computed using the Student's *t* test. An adjusted *p* value of 0.05 was used as the significance threshold.

shown in Figure 4B. We evaluated the impacts of fibrous stripes on pulmonary functions between the FIB and NFIB patient groups but did not observe statistically significant differences (Table S7).

Serum SFTPB and ATR are associated with fibrous-stripe-based pulmonary long COVID

We further applied temporal serum proteomics (SP2, Figure S4A) to compare the host responses in patients with FIB and NFIB during hospitalization (T1 and T3) and revisit (R1). Two groups of non-COVID individuals without (Ctrl1, *N* = 15) or with (Ctrl2, *N* = 14) fibrous stripes were further enrolled as control groups (Table S1). Among the 788 quantified proteins, SFTPB, HIST2H2BE, ATR, APOB, and DPT were differentially expressed between the FIB and NFIB groups (Figure 4C). Interestingly, the expression of serum pulmonary surfactant-associated protein B (SFTPB) was significantly up-regulated in the NFIB group compared to the FIB group at T1, and it also exceeded the expression in the Ctrl1 group (Figure 4D). This suggests that the up-regulation of SFTPB in the NFIB group was COVID-19 specific. SFTPB promotes alveolar stability and has been reported to be negatively associated with cystic fibrosis.[17] Its expression in hamster lung tissue post-SARS-CoV-2 infection increased at the recovery stage compared to the acute phase, suggesting its protective role in viral infection.[18] Conversely, the level of serum ataxia telangiectasia and Rad3-related protein (ATR), gradually increased from T1 to R1 (Figure 4E). ATR expression was suppressed in the 1-year revisited patients (FIB and NFIB) compared to that in the control groups (Ctrl1 and Ctrl2) at T1. However, at R1, ATR expression in the FIB group was significantly higher than that in the NFIB and Ctrl2 groups. Serum ATR is activated in COVID-19 owing to DNA damage, which mediates cell death and inflammation.[19] Monitoring the expression of serum SFTPB and ATR might assist in evaluating pulmonary epithelial injuries.

Emerging pulmonary fibrous stripes at two-year revisit

Three patients exhibited newly emerging fibrous stripes at R2 (R008, R046, and R130), all under the pleura and in the lower lobes (Figure 5). Two of these patients (R008 and R130) were over 50 years old and had severe symptoms during the disease. Chest CT images of the three patients at T1 showed large scales of GGO and consolidation shadows under the pleura of both lungs, involving at least three lobes, fuzzy edges, and an area covering 40%–60% of the entire lung. At T3, these lesions were partially absorbed, and the PO/GGO lesions replaced several consolidation lesions. At R1, the consolidation lesions were completely absorbed in these patients. The GGO shadows were also significantly absorbed in these patients, with a residual lesion area ranging from 5% to 30% of the whole lung and involving two to five lobes.
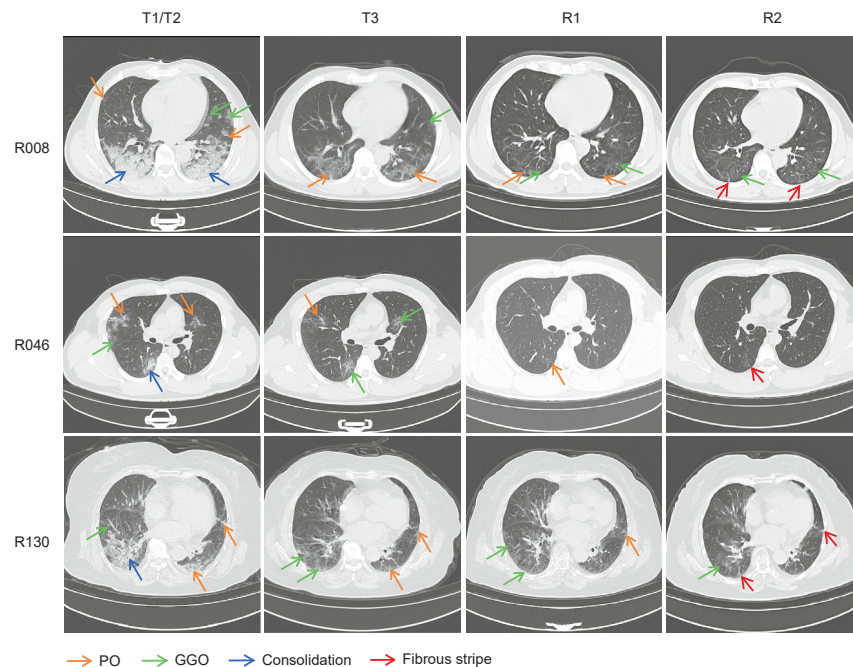
Figure 5. Emerging cases of pulmonary long COVID at two years
Longitudinal chest CT images of three patients with emerged fibrous-stripe-based pulmonary long COVID at R2.

Of the remaining PO/GGO lesions at R1, five had turned into fibrous stripe lesions at R2. GGO persisted in patients R008 and R130 throughout the 2 years of monitoring. These observations indicated that long-standing lesions may evolve into fibrous stripes.

Persistent dysregulation of host responses in renal long COVID

Renal dysfunction is a frequently reported symptom of COVID-19.[20] After excluding four patients younger than 18 years of age or those with underlying renal disease to reduce confounders (STAR methods), the enrolled patients in our study exhibited significantly increased eGFR at R1 than at T1 (Figure 6A). However, most of the patients experienced a decrease in eGFR at R2, suggesting the possibility of delayed renal injury. Notably, eight patients showed abnormal eGFRs at both T1 and R1, whereas only one patient recovered at R2 (Table S8), suggesting that their renal dysfunction persisted for years. Supportively, previous studies have reported renal abnormalities in patients with COVID-19, such as eGFR down-regulation upon admission[21] and at nearly 1 year after symptom onset.[22] These eight patients were classified as the abnormal renal function group, while the other patients with normal eGFRs at R1 were classified as the normal renal function group (N = 61, Figure 6A).

Thereafter, we compared the urine proteomes (UP2, Figure S4A) between patients with normal and abnormal renal function. Of the 4,161 quantified urinary proteins, 373 and 151 were differentially expressed at T1 and R1, respectively, suggesting a persistent dysregulation of the host response in patients with abnormal renal function. Among the ten most significantly down-regulated urine proteins in the abnormal renal function group at T1, six were associated with the endosomal sorting complexes required for transport (ESCRT) machinery (Figure 6B). The most significant differentially expressed proteins (DEPs) in the abnormal renal function group at R1 included the down-regulated proenkephalin (PENK), a molecular biomarker of kidney dysfunction,[23] and the up-regulated MASP1, proposed as a COVID-19 risk factor.[24] Notably, 23 DEPs overlapped at T1 and R1, nine of which could be attributed to the impairment of the renal glomerulus, proximal tubule, epithelium, and cortex (Figure 6C). These nine proteins were down-regulated in the abnormal renal function group at both T1 and R1, suggesting that diverse renal injuries may persist for up to 1 year after COVID-19 infection.

Five patients exhibited newly emerging renal abnormalities 2 years after COVID-19, with a sharp decrease in eGFR levels from R1 to R2 (Figure 6D). All the patients were above 50 years old. Four of them had abnormal renal function at their first sampling upon admission, and among them, three had severe COVID-19. These observations suggest the possibility of delayed renal long COVID symptoms.

DISCUSSION

We leveraged proteomic, metabolomic, clinical, and imaging data obtained from hospitalized patients previously enrolled during the first COVID-19 pandemic in 2020. This analysis focused on 89 patients who were subsequently enrolled for follow-up visits at 1- and 2-year for the assessment of pulmonary and renal long COVID-19, with further clinical assessments, multi-omics assays, and chest CT imaging.
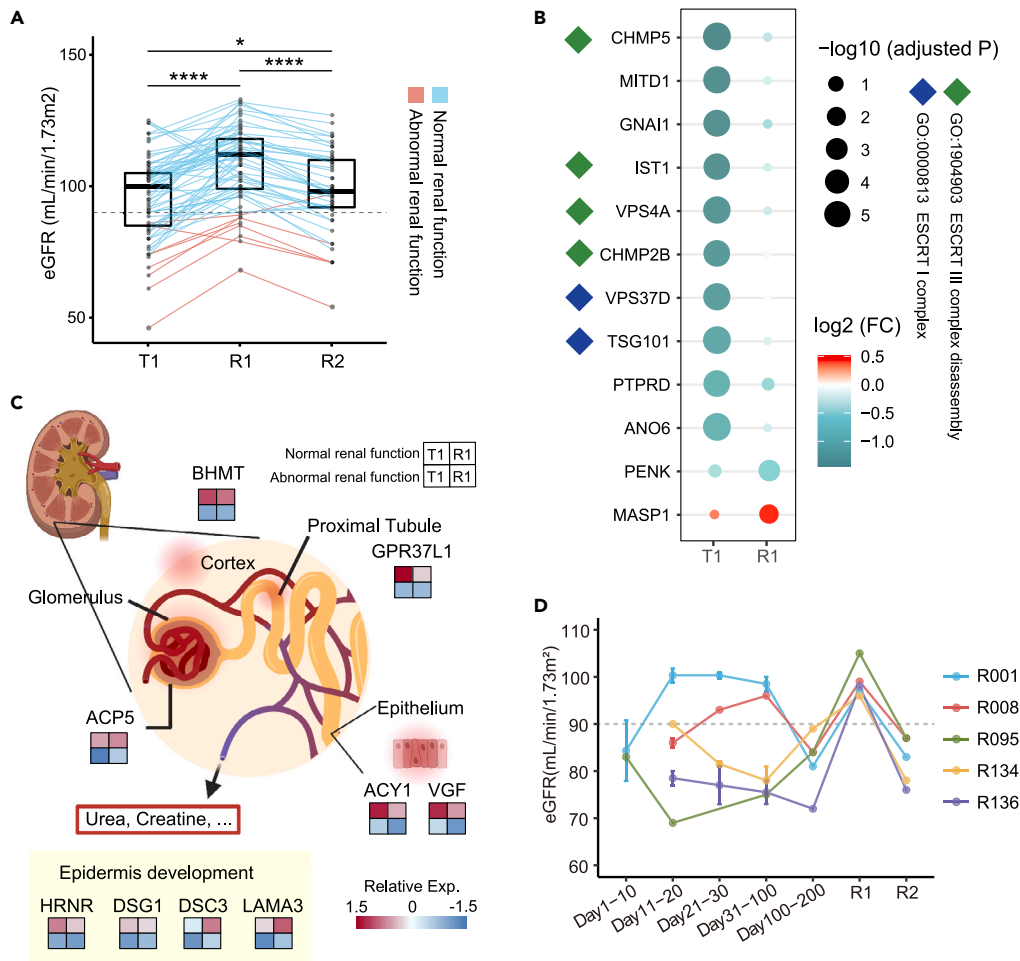
Figure 6. Renal long COVID in the recovered patients

(A) Classification of the patients with normal and abnormal renal functions based on their temporal eGFR expression. Data are represented as median (interquartile range). *p* values were calculated using two-sided Wilcoxon rank-sum tests for each pair of time points. The signs of significance were described in *p* values, *, *p* < 0.05; **, *p* < 0.01; ***, *p* < 0.005; ****, *p* < 0.001.

(B) Twelve selected urine DEPs at T1 and R1.

(C) Working model of the kidney injuries derived from the shared DEPs between T1 and R1.

(D) Longitudinal eGFR levels of five patients with emerged renal long COVID at R2. DEP, differentially expressed protein.

Several other studies have investigated the hypothetical pathogenesis of long COVID, including immune disorders, microbiota destruction, and autoimmunity.[4] Proteomics analysis showed that tissue damage and repair-related inflammatory mediators increased in the moderate and severe cognitive impairment group at 5 months.[25] However, these studies did not explicitly assess pulmonary and renal long COVID, and the monitoring periods were limited to within 1 year.

Both the lungs and kidneys express ACE2, a canonical SARS-CoV-2 receptor. Indeed, both pulmonary and renal injuries have been observed in patients with COVID-19 patients.[20,26] Moreover, our data revealed persistent abnormalities in these two organs. We developed a machine learning-based model to predict long COVID that incorporates serum and urinary molecular markers, and clinical indices, providing a flexible option for assessing the possibility of developing pulmonary and renal long COVID in patients with COVID-19. Notably, one patient with long COVID was incorrectly classified as recovered owing to his medical history of diabetes mellitus. Diabetes has been reported to play a role in COVID-19.[27] Our findings suggest the need to actively monitor patients with underlying diseases.

Furthermore, the prediction model established in this study uses data from the acute phase of the disease to predict pulmonary and renal long COVID at 1-year, which holds particular clinical value. A total of 20 serum protein markers with biological relevance were validated in the model. However, further verification and studies on their regulatory mechanisms are required. Additionally, the serum protein marker panel has the potential to predict pulmonary and renal long COVID in the longer terms.

Fibrous stripes are common CT manifestations in COVID-19.[28] A recent study showed that fibrous stripes, observed in 51.2% of convalescent patients with COVID-19, are the most predominant radiological abnormality 1 year after discharge.[29] Additionally, a long-term follow-up

of 11 convalescent patients with severe acute respiratory syndrome (SARS) showed that lung fibrosis manifestations, such as interlobular or intralobular thickening and subpleural lines, were observed 7 years after infection,[30] suggesting that early fibrous stripe lesions may require long-term follow-up.

Our study might provide indications for medical interventions for pulmonary long COVID. Two serum proteins, SFTPB and ATR, were strongly associated with fibrous stripe-based pulmonary long COVID. Particularly, the up-regulation of SFTPB could be essential in regulating lung protection. *SFTPB* encodes the lung-associated surfactant protein B (SP-B), which is necessary for lung function. It mainly promotes the formation and maintenance of phospholipid-rich membranes at the alveolar gas-liquid interface. Researchers speculate that *SFTPB* variants can affect the processing, secretion, and folding of SP-B in certain diseases.[31] SFTPB participates in immune regulation and plays an essential role in the defense and stability of the host airway. SFTPB deficiency is associated with various lung diseases, such as fatal respiratory failure, idiopathic pulmonary fibrosis, allergic pneumonia, chronic obstructive pulmonary disease, and acute respiratory distress syndrome.[31,32] Therefore, it can be inferred that SFTPB might be a potential biomarker for early and later monitoring of pulmonary long COVID. In contrast, persistent GGO lesions may turn into fibrous stripes, as observed in the newly emerged pulmonary long COVID cases in our cohort. A previous study on SARS-CoV attributes this transition to residual inflammation.[33] Our data suggest that persistent GGO lesions during revisits in convalescent patients with COVID-19 should be actively monitored for long-term lung health.

Our study uncovered two potential characteristics of renal long COVID. First, perturbation of ESCRT function from disease onset contributes to the 1-year renal long COVID. Notably, the most significant urine protein marker in the machine learning model, PDCD6IP, is also a member of the ESCRT pathway, which functions in the abscission stage of cytokinesis.[34] Our previous study reported that decreased urinary ESCRT complex proteins might be associated with active SARS-CoV-2 replication and that regulated innate immunity and inflammatory activities contribute to the resultant renal injuries.[10] Second, we observed delayed renal injury in some patients until the second year, which might have been a delayed COVID-19 symptom. A study on the progression of spontaneous kidney disease in mice showed that renal tubular injury preceded glomerular filtration function, leading to a delayed eGFR decrease.[35] The delayed reduction in eGFR at R2 stage observed in our study might have been due to similar processes. Our data suggest that elderly patients with severe COVID-19, particularly those with dysfunctional kidneys during hospitalization, should be actively monitored for renal function.

Conclusion

This study provides an integrated molecular landscape resource based on patients who were available for the 1-year and 2-year follow-up visits after hospitalization for COVID-19. Combining clinical data with multi-omics profiles, the machine learning model could predict pulmonary and renal long COVID at 1 year based on one month of data since the onset of the disease. Serum proteomics highlighted pulmonary fibrous stripes as the most characteristic symptom of pulmonary long COVID, which might be associated with the continuous down-regulation of SFTPB and the up-regulation of ATR. The proportion of renal long COVID increased with follow-up time, possibly due to the inhibition of urinary protein expression during the disease.

Limitations of the study

This study had several limitations. Because this was a single-center study with a relatively small patient cohort, it was possibly subject to demographic and experimental biases. Only 31.9% of the patients consented for revisits both at R1 and R2 owing to the extended distance from our clinical center. Additionally, owing to the relatively small sample size, the covariables explored in the study were not corrected for multivariability to exclude confounders that might influence progression into chronicity, such as the severity of acute illness and age. The small sample size for predictive models might lead to variability in the data, including multi-omics data collected from the original 144 patients enrolled. However, we included two independent cohorts with different patients and samples to consolidate the robustness of our ensembled model. Hepatic dysfunctions during and after COVID-19 are multifactorial[36]; therefore, they were not analyzed in detail in this study. Furthermore, although we had generated a resource with two years of patient data, we only used the multi-omics and clinical data from the index hospitalization (initial enrollment cohorts of Shen, Bi, Zhang) and 1-year pulmonary and renal long COVID outcomes to create our prediction models. Furthermore, we used serum and urine samples from 1-year revisit patients to explore the molecules related to pulmonary and renal long COVID. However, the second-year data were only used for clinical manifestations, considering the involvement of excessive contributing factors, such as possible influenza virus infections, and insufficient patient numbers for machine learning. Future validation of the model using independent cohorts with more stringent enrollment criteria is required. In addition, this study combined the primary outcomes of pulmonary or renal abnormalities to denote long COVID for the models; however, we did not explore the correlation between pulmonary and renal long COVID. This study only focused on the original SARS-CoV-2 strain. The same analysis could be applied to other strains or even other infectious diseases; however, this was beyond the scope of the current study.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110344.

AUTHOR CONTRIBUTIONS

B.S. and T.G. designed the project. H.C., J.L., J.X., B.S., and X.W. supervised the project. J.W., Y. Zhang, K.Z., H.C., D.W., and G.Z. collected the samples. J.W., Y. Zhang, K.Z., X.Z., H.C., D.W., and M.P. collected the clinical data. X.W., H.Z., S.S., J.X., H.Y., and J.L. organized the clinical data. J.P. and Y.F. performed CT assessment. J.W., Y. Zhang, K.Z., X.Z., and H.C. performed clinical data analyses. X.L., X.Y., S.L., W.L., X.C., Q.S, C.W., H.L., C.Z., S.Q., Z.Kang and Z. Kong conducted multi-omics analyses. X.L. collected the MS-based multi-omics data. X.L., R.S., and M.L. performed biological interpretation. X.L. and Y.H. built the prediction model. J.W., X.L., Y.S., and Y.X. designed the figures. J.W., X.L., Y. Zhang, X.Z., K.Z., T.G., and Y. Zhu wrote the manuscript with inputs from all co-authors.

DECLARATION OF INTERESTS

S.C., X.Y., W.L., Q.Z., and C.W. are employees of Westlake Omics (Hangzhou) Biotechnology Co. H.L., C.Z., S.Q., Z.Kang, and Z.Kong are employees of DIAN Diagnostics. T.G. and Y. Z. are shareholders of Westlake Omics (Hangzhou) Biotechnology Co.

REFERENCES

1. WHO (2022). WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/.

2. Huang, L., Yao, Q., Gu, X., Wang, Q., Ren, L., Wang, Y., Hu, P., Guo, L., Liu, M., Xu, J., et al. (2021). 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. Lancet (London, England) 398, 747–758. https://doi.org/10.1016/s0140-6736(21)01755-4.

3. Crook, H., Raza, S., Nowell, J., Young, M., and Edison, P. (2021). Long covid-mechanisms, risk factors, and management. BMJ 374, n1648. https://doi.org/10.1136/bmj.n1648.

4. Davis, H.E., McCorkell, L., Vogel, J.M., and Topol, E.J. (2023). Long COVID: major findings, mechanisms and recommendations. Nat. Rev. Microbiol. 21, 133–146. https://doi.org/10.1038/s41579-022-00846-2.

5. Klein, J., Wood, J., Jaycox, J., Lu, P., Dhodapkar, R.M., Gehlhausen, J.R., Tabachnikova, A., Tabacof, L., Malik, A.A., Kamath, K., et al. (2022). Distinguishing features of Long COVID identified through immune profiling. medRxiv 623, 139–148.

https://doi.org/10.1101/2022.08.09.22278592.

6. Mullard, A. (2021). Long COVID's long R&D agenda. Nat. Rev. Drug Discov. 20, 329–331. https://doi.org/10.1038/d41573-021-00069-9.

7. den Hartog, G., Vos, E.R.A., van den Hoogen, L.L., van Boven, M., Schepp, R.M., Smits, G., van Vliet, J., Woudstra, L., Wijmenga-Monsuur, A.J., van Hagen, C.C.E., et al. (2021). Persistence of Antibodies to Severe Acute Respiratory Syndrome Coronavirus 2 in Relation to Symptoms in a Nationwide Prospective Study. Clin. Infect. Dis. 73, 2155–2162. https://doi.org/10.1093/cid/ciab172.

8. Wu, X., Liu, X., Zhou, Y., Yu, H., Li, R., Zhan, Q., Ni, F., Fang, S., Lu, Y., Ding, X., et al. (2021). 3-month, 6-month, 9-month, and 12-month respiratory outcomes in patients following COVID-19-related hospitalisation: a prospective study. Lancet Respir. Med. 9, 747–754. https://doi.org/10.1016/s2213-2600(21)00174-0.

9. Huang, C., Huang, L., Wang, Y., Li, X., Ren, L., Gu, X., Kang, L., Guo, L., Liu, M., Zhou, X., et al. (2021). 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. Lancet 397, 220–232. https://doi.org/10.1016/S0140-6736(20)32656-8.

10. Bi, X., Liu, W., Ding, X., Liang, S., Zheng, Y., Zhu, X., Quan, S., Yi, X., Xiang, N., Du, J., et al. (2022). Proteomic and metabolomic profiling of urine uncovers immune responses in patients with COVID-19. Cell Rep. 38, 110271. https://doi.org/10.1016/j.celrep.2021.110271.

11. Shen, B., Yi, X., Sun, Y., Bi, X., Du, J., Zhang, C., Quan, S., Zhang, F., Sun, R., Qian, L., et al. (2020). Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. Cell 182, 59–72.e15. https://doi.org/10.1016/j.cell.2020.05.032.

12. Zhang, Y., Cai, X., Ge, W., Wang, D., Zhu, G., Qian, L., Xiang, N., Yue, L., Liang, S., Zhang, F., et al. (2022). Potential Use of Serum Proteomics for Monitoring COVID-19 Progression to Complement RT-PCR Detection. J. Proteome Res. 21, 90–100. https://doi.org/10.1021/acs.jproteome.1c00525.

13. Xiang, F., Wang, X., He, X., Peng, Z., Yang, B., Zhang, J., Zhou, Q., Ye, H., Ma, Y., Li, H., et al. (2020). Antibody Detection and Dynamic Characteristics in Patients With Coronavirus Disease 2019. Clin. Infect. Dis. 71, 1930–1934. https://doi.org/10.1093/cid/ciaa461.

14. Bonaventura, A., Vecchié, A., Dagna, L., Martinod, K., Dixon, D.L., Van Tassell, B.W., Dentali, F., Montecucco, F., Massberg, S., Levi, M., and Abbate, A. (2021). Endothelial dysfunction and immunothrombosis as key pathogenic mechanisms in COVID-19. Nat. Rev. Immunol. 21, 319–329. https://doi.org/10.1038/s41577-021-00536-9.

15. Wang, B., Yan, X., Chen, F., Yang, A., Lu, Y., and Wu, Y. (2019). Plasma kallikrein contributes to ambient particulate matter-induced lung injury. Biochem. Biophys. Res. Commun. 518, 409–415. https://doi.org/10.1016/j.bbrc.2019.07.060.

16. Zhi, H., Ji, X., Zhao, Z., Liang, H., Zhong, S., Luo, Y., Zhong, M., Zhan, C., Gao, Y., Deng, X., et al. (2022). Risk factors for impaired pulmonary diffusion function in convalescent COVID-19 patients: A systematic review and meta-analysis. EClinicalMedicine 49, 101473.

https://doi.org/10.1016/j.eclinm.2022.101473.

17. Lin, Z., Thorenoor, N., Wu, R., DiAngelo, S.L., Ye, M., Thomas, N.J., Liao, X., Lin, T.R., Warren, S., and Floros, J. (2018). Genetic Association of Pulmonary Surfactant Protein Genes, SFTPA1, SFTPA2, SFTPB, SFTPC, and SFTPD With Cystic Fibrosis. Front. Immunol. 9, 2256. https://doi.org/10.3389/fimmu.2018.02256.

18. Suresh, V., Mohanty, V., Avula, K., Ghosh, A., Singh, B., Reddy, R.K., Parida, D., Suryawanshi, A.R., Raghav, S.K., Chattopadhyay, S., et al. (2021). Quantitative proteomics of hamster lung tissues infected with SARS-CoV-2 reveal host factors having implication in the disease pathogenesis and severity. FASEB J. 35, e21713. https://doi.org/10.1096/fj.202100431R.

19. Riebeling, T., Jamal, K., Wilson, R., Kolbrink, B., von Samson-Himmelstjerna, F.A., Moerke, C., Ramos Garcia, L., Dahlke, E., Michels, F., Lühder, F., et al. (2021). Primidone blocks RIPK1-driven cell death and inflammation. Cell Death Differ. 28, 1610–1626. https://doi.org/10.1038/s41418-020-00690-y.

20. Ronco, C., Reis, T., and Husain-Syed, F. (2020). Management of acute kidney injury in patients with COVID-19. Lancet Respir. Med. 8, 738–742. https://doi.org/10.1016/s2213-2600(20)30229-0.

21. Cheng, Y., Luo, R., Wang, K., Zhang, M., Wang, Z., Dong, L., Li, J., Yao, Y., Ge, S., and Xu, G. (2020). Kidney disease is associated with in-hospital death of patients with COVID-19. Kidney Int. 97, 829–838. https://doi.org/10.1016/j.kint.2020.03.005.

22. Gu, X., Huang, L., Cui, D., Wang, Y., Wang, Y., Xu, J., Shang, L., Fan, G., and Cao, B. (2022). Association of acute kidney injury with 1-year outcome of kidney function in hospital survivors with COVID-19: A cohort study. EBioMedicine 76, 103817. https://doi.org/10.1016/j.ebiom.2022.103817.

23. Khorashadi, M., Beunders, R., Pickkers, P., and Legrand, M. (2020). Proenkephalin: A New Biomarker for Glomerular Filtration Rate and Acute Kidney Injury. Nephron 144, 655–661. https://doi.org/10.1159/000509352.

24. Bumiller-Bini, V., de Freitas Oliveira-Toré, C., Carvalho, T.M., Kretzschmar, G.C., Gonçalves, L.B., Alencar, N.d.M., Gasparetto Filho, M.A., Beltrame, M.H., and Winter Boldt, A.B. (2021). MASPs at the crossroad between the complement and the coagulation cascades - the case for COVID-19. Genet. Mol. Biol. 44, e20200199. https://doi.org/10.1590/1678-4685-gmb-2020-0199.

25. Evans, R.A., Leavy, O., Richardson, M., and Brightling, C.E. (2021). Clinical Characteristics with Inflammation Profiling of Long-COVID and Association with One-Year Recovery Following Hospitalisation in the UK: A Prospective Observational Study.

26. Monteil, V., Kwon, H., Prado, P., Hagelkrüys, A., Wimmer, R.A., Stahl, M., Leopoldi, A., Garreta, E., Hurtado Del Pozo, C., Prosper, F., et al. (2020). Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues Using Clinical-Grade Soluble Human ACE2. Cell 181, 905–913.e7. https://doi.org/10.1016/j.cell.2020.04.004.

27. Tong, L., Xiao, X., Li, M., Fang, S., Ma, E., Yu, X., Zhu, Y., Wu, C., Tian, D., Yang, F., et al. (2022). A glucose-like metabolite deficient in diabetes inhibits cellular entry of SARS-CoV-2. Nat. Metab. 4, 547–558. https://doi.org/10.1038/s42255-022-00567-z.

28. Niu, R., Ye, S., Li, Y., Ma, H., Xie, X., Hu, S., Huang, X., Ou, Y., and Chen, J. (2021). Chest CT features associated with the clinical characteristics of patients with COVID-19 pneumonia. Ann. Med. 53, 169–180. https://doi.org/10.1080/07853890.2020.1851044.

29. Liu, T., Wu, D., Yan, W., Wang, X., Zhang, X., Ma, K., Chen, H., Zeng, Z., Qin, Y., Wang, H., et al. (2022). Twelve-Month Systemic Consequences of Coronavirus Disease 2019 (COVID-19) in Patients Discharged From Hospital: A Prospective Cohort Study in Wuhan, China. Clin. Infect. Dis. 74, 1953–1965. https://doi.org/10.1093/cid/ciab703.

30. Wu, X., Dong, D., and Ma, D. (2016). Thin-Section Computed Tomography Manifestations During Convalescence and Long-Term Follow-Up of Patients with Severe Acute Respiratory Syndrome (SARS). Med. Sci. Monit. 22, 2793–2799. https://doi.org/10.12659/msm.896985.

31. Fishchuk, L., Rossokha, Z., Pokhylko, V., Cherniavska, Y., Popova, O., Vershyhora, V., Kovtun, S., and Gorovenko, N. (2023). SFTPB (rs11130866) and NR3C1 (rs41423247) gene variants as potential clinical biomarkers for personalized treatment strategy selection in patients with severe COVID-19 pneumonia. Respir. Investig. 61, 103–109. https://doi.org/10.1016/j.resinv.2022.10.008.

32. Kang, M.H., van Lieshout, L.P., Xu, L., Domm, J.M., Vadivel, A., Renesme, L., Mühlfeld, C., Hurskainen, M., Mižíková, I., Pei, Y., et al. (2020). A lung tropic AAV vector improves survival in a mouse model of surfactant B deficiency. Nat. Commun. 11, 3929. https://doi.org/10.1038/s41467-020-17577-8.

33. Hsu, H.H., Tzao, C., Wu, C.P., Chang, W.C., Tsai, C.L., Tung, H.J., and Chen, C.Y. (2004). Correlation of high-resolution CT, symptoms, and pulmonary function in patients during recovery from severe acute respiratory syndrome. Chest 126, 149–158. https://doi.org/10.1378/chest.126.1.149.

34. Murrow, L., Malhotra, R., and Debnath, J. (2015). ATG12-ATG3 interacts with Alix to promote basal autophagic flux and late endosome function. Nat. Cell Biol. 17, 300–310. https://doi.org/10.1038/ncb3112.

35. Kasztan, M., Fox, B.M., Lebensburger, J.D., Hyndman, K.A., Speed, J.S., Pollock, J.S., and Pollock, D.M. (2019). Hyperfiltration predicts long-term renal outcomes in humanized sickle cell mice. Blood Adv. 3, 1460–1475. https://doi.org/10.1182/bloodadvances.2018028878.

36. Jothimani, D., Venugopal, R., Abedin, M.F., Kaliamoorthy, I., and Rela, M. (2020). COVID-19 and the liver. J. Hepatol. 73, 1231–1240. https://doi.org/10.1016/j.jhep.2020.06.006.

37. Kroemer, G., Galassi, C., Zitvogel, L., and Galluzzi, L. (2022). Immunogenic cell stress and death. Nat. Immunol. 23, 487–500. https://doi.org/10.1038/s41590-022-01132-2.

38. Xie, M., Cui, L., Liu, J., Wang, W., Li, J., and Xiao, W. (2020). Impacts of Different Spirometry Reference Equations and Diagnostic Criteria on the Frequency of Airway Obstruction in Adult People of North China. Int. J. Chron. Obstruct. Pulmon. Dis. 15, 651–659. https://doi.org/10.2147/copd.S232863.

39. Tai, D.Y., Wang, Y.T., Lou, J., Wang, W.Y., Mak, K.H., and Cheng, H.K. (1996). Lungs in thalassaemia major patients receiving regular transfusion. Eur. Respir. J. 9, 1389–1394. https://doi.org/10.1183/09031936.96.09071389.

40. Quanjer, P.H., and Weiner, D.J. (2014). Interpretative consequences of adopting the Global Lungs 2012 reference equations for spirometry for children and adolescents. Pediatr. Pulmonol. *49*, 118–125. https://doi.org/10.1002/ppul.22876.

41. Xiao, D., Chen, Z., Wu, S., Huang, K., Xu, J., Yang, L., Xu, Y., Zhang, X., Bai, C., Kang, J., et al. (2020). Prevalence and risk factors of small airway dysfunction, and association with smoking, in China: findings from a national cross-sectional study. Lancet Respir. Med. *8*, 1081–1093. https://doi.org/10.1016/s2213-2600(20)30155-7.

42. Xie, L., Liu, Y., Fan, B., Xiao, Y., Tian, Q., Chen, L., Zhao, H., and Chen, W. (2005). Dynamic changes of serum SARS-coronavirus IgG, pulmonary function and radiography in patients recovering from SARS after hospital discharge. Respir. Res. *6*, 5. https://doi.org/10.1186/1465-9921-6-5.

43. Inker, L.A., Astor, B.C., Fox, C.H., Isakova, T., Lash, J.P., Peralta, C.A., Kurella Tamura, M., and Feldman, H.I. (2014). KDOQI US commentary on the 2012 KDIGO clinical practice guideline for the evaluation and management of CKD. Am. J. Kidney Dis. *63*, 713–735. https://doi.org/10.1053/j.ajkd.2014.01.416.

44. Malhotra, K., Katsanos, A.H., Goyal, N., Tayal, A., Gensicke, H., Mitsias, P.D., De Marchis, G.M., Berge, E., Alexandrov, A.W., Alexandrov, A.V., and Tsivgoulis, G. (2020). Intravenous thrombolysis in patients with chronic kidney disease: A systematic review and meta-analysis. Neurology *95*, e121–e130. https://doi.org/10.1212/wnl.0000000000009756.

45. Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (Association for Computing Machinery).

46. MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., and MacCoss, M.J. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics *26*, 966–968. https://doi.org/10.1093/bioinformatics/btq054.

47. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523. https://doi.org/10.1038/s41467-019-09234-6.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Triethylammonium bicarbonate buffer (TEAB)	Sigma-Aldrich	Cat #T7408
Urea	Sigma-Aldrich	Cat #U1250
Tris (2-carboxyethyl) phosphine (TCEP)	Adamas-beta	Cat # 61820E
Iodoacetamide (IAA)	Sigma-Aldrich	Cat #I6125
Trypsin	Hualishi Tech	Cat # HLS TRY001C
Lys-C	Hualishi Tech	Cat # HLS LYS001C
Trifluoroacetic acid (TFA)	Thermo Fisher Scientific	Cat # 85183
Water	Thermo Fisher Scientific	Cat #W6-4
Acetonitrile	Thermo Fisher Scientific	Cat # A955-4
Formic acid (FA)	Thermo Fisher Scientific	Cat # A117-50
Ammonium hydroxide solution	Sigma-Aldrich	Cat # 221228
Methanol	Sigma-Aldrich	Cat # 3486
Critical commercial assays		
Cytokine and lymphocyte subsets kit during hospitalization	Becton, Dickinson and Company	Cat # 33593
Cytokine and lymphocyte subsets kit at one-year revisit	Becton, Dickinson and Company	Cat #92766
Cytokine and lymphocyte subsets kit at two-year revisit	Becton, Dickinson and Company	Cat #23747
Urea kit during hospitalization	Beckman Coulter	Cat # AUZ6735
Urea kit at one-year revisit	Beckman Coulter	Cat #AUZ8361
Urea kit at two-year revisit	Beckman Coulter	Cat #AUZ9369
Creatinine kit during hospitalization	Beckman Coulter	Cat # 2587
Creatinine kit at one-year revisit	Ausbio	Cat #191012
Creatinine kit at two-year revisit	Ausbio	Cat #210726
Anti streptolysin O kit during hospitalization	Beckman Coulter	Cat # 2586
Anti streptolysin O kit at one-year revisit	Beckman Coulter	Cat #2591
Anti streptolysin O kit at two-year revisit	Beckman Coulter	Cat #2602
Potassium kit during hospitalization	Beckman Coulter	Cat # AUZ6952
Potassium kit at one-year revisit	Beckman Coulter	Cat #AUZ8696
Potassium kit at two-year revisit	Beckman Coulter	Cat #AUZ9434
Activated partial thromboplastin time kit during hospitalization	Sysmex	Cat # 556966
Activated partial thromboplastin time kit at one-year revisit	Stago	Cat # 256815
Activated partial thromboplastin time kit at two-year revisit	Stago	Cat # 259331
Complete blood count kit during hospitalization	Sysmex	Cat # FFD-A8099; Cat # FFS-A9047; Cat # SE-G8002; Cat # PK-G9250; Cat #CL-50- A8001; Cat #FBA- R7057; Cat #SLS-A9002

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Complete blood count at one-year revisit	Sysmex	Cat #CELLPACK DST-A0005; Cat #CELLPACK DFL-A0007; Cat #SULFOLYSER-A0012; Cat #Lysercell WPC-A0007; Cat #Lysercell WNR-A0027; Cat #Lysercell WDF-A0027; Cat #Fluorocell WPC-A0013; Cat #Fluorocell WNR-A0001; Cat #Fluoroc ell WDF-A0054.
Complete blood count at two-year revisit	Sysmex	Cat #CELLPACK DST-A1001; Cat #CELLPACK DFL-A1001; Cat #SULFOLYSER-A1011; Cat #Lysercell WPC-A0012; Cat #Lysercell WNR-A1032; Cat #Lysercell WDF-A1018; Cat #Fluorocell WPC-A1005; Cat #Fluorocell WNR-A1041; Cat #Fluoroc ell WDF-A1058; Cat #DCL-300A-G1424
Apolipoprotein A1 kit during hospitalization	Beckman Coulter	Cat # AA0183
Apolipoprotein A1 kit at one-year revisit	Leadman	Cat #20091311
Apolipoprotein A1 kit at two-year revisit	Leadman	Cat #21101217
TMTpro 16plex reagents	Thermo Fisher Scientific	Cat # A44520
High SelectTM Top-14 Abundant Protein Depletion Mini Spin Columns	Thermo Fisher Scientific	Cat #A36370
SOLAμ	Thermo Fisher Scientific	Cat # 62209-001
Pierce™ C18 Spin Columns	Thermo Fisher Scientific	Cat # 89870
Deposited data		
Mass spectrometry data	https://www.iprox.cn//page/project.html?id=IPX0003240000	N/A
Software and algorithms		
R 3.6.3	http://www.rstudios.co	N/A
Machine learning codes	https://github.com/guomics-lab/CVDSBR	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Bo Shen (shenb@enzemed.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data are available in the manuscript or the supplementary material. The MS files of proteomic datasets SP2, SP3, UP2, and the validation dataset were deposited at ProteomeXchange Consortium. Project ID: IPX0003240000, (https://www.iprox.cn//page/project.html?id=IPX0003240000) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. The MS files for other previously published multi-omics datasets[10–12] and deposited in the iProx Consortium. Project ID: IPX0002106000 (SP1), IPX0002171000 (SM1), IPX0002365000 (SP5, UP1, and UP3), and SP4 (IPX0002374000).

- All original code of machine learning model generation has been deposited at GitHub and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

144 COVID-19 patients had been previously enrolled in other studies of our team during their index hospitalization during acute illness, 108 were diagnosed as non-severe and 36 as severe, and patient R124 was further diagnosed as a critically ill patient.[10–12] All patients were infected with the original SARS-CoV-2 strain; none were re-infected during the manifestation.

For this study, we contacted the original 144 patients previously enrolled for recruitment for 1- and 2-year follow-up visits for assessment of long COVID-19. Patients were excluded from the one-year follow-up when one of the following conditions was fulfilled: 1) refusal of participation; 2) unavailability of contact; 3) living outside Taizhou. After exclusion, eighty-nine patients were enrolled in this study, including 47 males and 42 females with a mean age of 47.0 (sd ± 14.1) years, the detail information was shown in Table S1.

Among them 89 patients, seventy-three were enrolled during day 379–389 (median, 385) since disease onset (Table S1). After excluding two cases who had a self-reported medical history of chronic renal diseases, two cases under 18 years old, two who did not have available pulmonary function tests, and one patient whose chest CT was examined one month before R1, 66 patients were finally enrolled for modeling. Among them, 75% of the cohort (50/66) was selected for the training set (chosen randomly) and held out 25% (16/66) for the internal test set. To verify the model, an independent cohort of 15 patients, including 7 patients who participated in R1 revisit but were not included in the modeling group, as well as 8 patients who did not participate in R1 revisit and with pulmonary and renal long COVID outcomes but without self-reported symptoms at R1, were enrolled as an external test set. Fifty-eight patients were enrolled in the two-year follow-up during day 745–752 (median, 749) since disease onset (Table S1). After excluding one patient, R072, without self-reporting information, the remaining 57 patients were included in the analysis. The inclusion details of each cohort were shown in Figure S1 and Table S4.

Forty non-COVID individuals who participated in the regular medical examination in Taizhou Public Health Center were enrolled as the control cohort, with gender, age, and BMI distribution equivalent to the one-year follow-up cohort (Table S1). Twenty-nine with chest CT examinations and serum proteomics sampling were grouped into Ctrl1 and Ctrl2 for pulmonary long COVID analysis.

Long COVID is defined as long-term effects from SARS-CoV-2 infection, including a wide range of new, returning, or ongoing health problems after COVID-19, according to the Centers for Disease Control and Prevention (CDC).[37] This study defined pulmonary long COVID as chest CT lesions in the COVID-19 convalescents, including patchy opacity, ground glass opacity, consolidation and fibrous stripe. Renal long COVID was defined as lasting abnormal kidney during hospitalization and COVID-19 convalescents, and hepatic long COVID was defined as abnormal liver functions in the COVID-19 convalescents.

This study was conducted following the principles of the Helsinki declaration. This study has been approved by the ethical committees of Taizhou Hospital of Zhejiang Province (Approve number: K20210218) and Westlake University and was also registered in the China Clinical Trial Registry with the registration number ChiCTR2100048440. Written informed consent regarding the samples and data collected during revisits was obtained from all revisited patients. Samples collected before revisits were previously processed into peptides and archived. Clinical data collected before revisits were previously archived as electronic hospital records. The IRB approval documents have covered all sample collection and data analysis processes.

METHOD DETAILS

Monitoring stages

For clinical indicators analyses and chest CT examinations, T1 was defined as the initial sampling within seven days from admission, while T3 was the last sampling within seven days before discharge. R1 was the sampling at the one-year follow-up. R2 was defined as the two-year follow-up. We further determined T2 in chest CT examination as samplings between T1 and T3.

For MS-based multi-omics sampling, T1 was the initial sampling from admission with a positive respiratory tract RNA result, while T3 was the last sampling before discharge with a negative respiratory tract RNA result. R1 was the sampling at the one-year follow-up.

Reagents and equipment

For a complete list of reagents and equipment used in this study, please refer to key resources table.

Flow cytometry analysis

The cytokine levels (IL-2, IL-4, IL-6, IL-10, TNF-α, and IFN-γ) and lymphocyte subsets were determined using the flow cytometer (FACS Canto TM II, BD, New Jersey, USA).

Cytokine detection

The cytokines were measured in plasma prepared from EDTA-K_2 anticoagulant. Six kinds of capture microspheres are available in the mixture of the kit (Nuode, Jiangxi, China). The capture microspheres coated with cytokines-specific antibodies explicitly bind to the cytokines in the patient samples. Then, the phycoerythrin (PE) marked antibody, the captured microspheres, and cytokines in the samples formed a double

antibody sandwich composite. The cytokine concentrations of the samples were measured by analyzing the fluorescence intensity of the double antibody sandwich composites.

Lymphocyte measurement

The Lymphocytes were measured in whole blood prepared from EDTA-K_2 anticoagulant. Fluorescent-labeled monoclonal antibodies bind to the lymphocyte surface antigens and lymphocyte subsets can be distinguished by flow cytometry according to the expression of cluster of differentiation (CD) molecules on the surface. The percentage of lymphocyte subsets was computed with BD FACSCanto Clinical Software and BD FACSDiva Software.

Scoring of chest CT images

Two senior radiologists independently read the chest CT images. Each image recorded the number of affected lung lobes and the presence of PO/GGO, consolidation, and fibrous stripe lesions. A CT score was calculated by summing up the scores for each type of lesion and the number of involved lobes. The existence of PO/GGO is considered +1, the existence of consolidation is considered +2, and the existence of fibrous stripe is considered +3.

Pulmonary function abnormality

The patient's pulmonary function was tested using the Power Cube-body ultrasonic pulmonary function meter (PowerCube-Body, GANSHORN, Niederlauer, Germany). The pulmonary function test was used to evaluate the patient's ventilation function, diffusion function, and residual air. The ventilation function test included the following indicators: FVC (Forced vital capacity); FEV1 (Forced expiratory volume in 1 s); PEF (Peak expiratory flow); FEF (Forced expiratory flow volume), *etc.* DLCO (Diffusion capacity of carbon monoxide), VA (Alveolar volume), and KCO (DLCO/VA) are parameters related to the diffusion function. The standard Helium dilution method was used to check for residual gas. Specifically, RV (residual gas volume), TLC (Total lung capacity), and RV/TLC were used to assess residual gas.

Patients with ventilation and/or diffusion dysfunction are diagnosed with abnormal pulmonary function. A ventilation dysfunction can be obstructive (FEV1/FVC <92% of the predicted value (pred))[38] or restrictive (TLC <80% pred, and for patients who did not have TLC records, FVC < LLN or <80% of pred accompany with FEV1/FVC ≥92% pred[39,40]). Patients with two of the three indicators in FEF50%, FEF75%, and FEF 25%–75% lower than the pred of 65% were diagnosed with small airway dysfunction.[41] Patients with more than one of the lung dysfunctions mentioned above were determined as ventilatory dysfunction. DLCO lower than 80% of the pred was considered a diffusion dysfunction.[42]

Renal function abnormality

The National Kidney Foundation–Kidney Disease Outcomes Quality Initiative guidelines defined chronic kidney disease as mild (eGFR 60–89 mL/min/1.73 m^2), moderate (eGFR 30–59 mL/min/1.73 m^2), or severe (eGFR 15–29 mL/min/1.73 m^2). In this study, an eGFR of 90 mL/min/1.73 m^2 was defined as the threshold, and an eGFR ≤89 mL/min/1.73 m^2 was chosen to determine an abnormal renal function.[43,44]

Patients under 18 years old or with underlying renal disease were excluded from the analysis.

Liver function abnormality

Patients with at least one of the following serum clinical indicators outside of the normal reference range were determined to have abnormal liver function: ALT (alanine aminotransferase), AST (aspartate aminotransferase), ALP (alkaline phosphatase), GGT (gamma-glutamine transferase), and total bilirubin.

Patients younger than 18 years or with underlying hepatic disease were excluded from the analysis.

Multi-omics data collection

Ten omics datasets were collected, seven of which were retrieved from three previously published studies, hereinafter labeled A, B, and C, respectively. Within them, study A[11] contains a serum proteomics (SP1) dataset and a serum metabolomics (SM1) dataset. Study B[12] contains a serum proteomics (SP4) dataset. Study C[10] includes a serum proteomics (SP5) dataset, two urine proteomics (UP1 and UP3) datasets, and a urine metabolomics (UM1) dataset. We included a targeted serum proteomics dataset for 22 serum proteins (SP3), covering 231 disease-stage samples from 66 patients. Additionally, we collected a serum proteomics (SP2) dataset and a urine proteomics (UP2) dataset, covering the multi-omics data during the revisit stage. A summary of their sample preparation strategies in Proteome analyses and Metabolome analyses was given below. An overview of their MS analysis workflow was provided in Figure S3A.

Proteome analyses

Blood samples from COVID-19 patients were centrifuged at 1500 g for 10 min to collect the serum. All serum samples were previously inactivated and sterilized at 56°C for 30 min. Next, 5 to 10 μL of serum was used for proteomic sample preparation.

The detailed experimental procedure of producing dataset SP2 is described below: the collected serum was denatured in 50 μL buffer composed of 8 M urea and100 mM triethylammonium bicarbonate (TEAB) at 32°C for 30 min. The proteins were reduced with 10 mM tris (2-carboxyethyl) phosphine (TCEP) for 30 min at 32°C and alkylated for 45 min with 40 mM iodoacetamide (IAA) in darkness at room

temperature. The protein extracts were diluted with 200 μL 100 mM TEAB, and digested with double-step trypsinization (Hualishi Tech. Ltd, Beijing, China), each step with an enzyme-to-substrate ratio of 1:20, at 32°C for 60 min. The reaction was stopped by adding 30 μL 10% tri-fluoroacetic acid (TFA). Digested peptides were cleaned-up with SOLAμ (Thermo Fisher Scientific, San Jose, USA) following the manufacturer's instructions. The clean peptides were labeled with the TMTpro 16plex label reagents (Thermo Scientific) according to the label set (Table S5). Channel 126 was set for pooled samples. The tandem mass tag (TMT)-labeled samples were separated using a nanoflow DIONEX UltiMate 3000 RSLCnano System (Thermo Fisher Scientific, San Jose, USA) with an XBridge Peptide BEH C18 column (300 Å, 5 μm × 4.6 mm × 250 mm) (Waters, Milford, MA, USA) (Gao et al., 2020) into 60 fractions and combined into 20 fractions. Peptides were sent to a Dionex Ultimate 3000 liquid chromatography system (Thermo Fisher Scientific) coupled with Q Exactive HF hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific). For each acquisition, peptides were analyzed using a 60 min liquid chromatography (LC) gradient (5%–28% of buffer B). Buffer A was 2% ACN, 98% H2O containing 0.1% FA, and buffer B was 98% ACN in water containing 0.1% FA. The m/z range of MS1 was 350-1,800 with the resolution at 60,000 (at 200 m/z), AGC target of 3e6, and maximum ion injection time (max IT) of 100 ms. Top 15 precursors were selected for MS/MS experiment, with a resolution at 45,000 (at 200 m/z), AGC target of 2e5, and max IT of 120 ms. The isolation window of selected precursor was 0.7 m/z. The MS data were then analyzed using Proteome Discoverer (Version 2.4.305, Thermo Fischer Scientific) and the protein database composed of the Homo sapiens SwissProt fasta database containing 20,365 reviewed protein sequences, downloaded from UniProtKB on April 14, 2020. Channel 126 was set as the denominator. Enzyme was set to trypsin with two missed cleavage tolerance. Static modifications were set to carbamidomethylation (+57.021464) of cysteine, TMTpro (+304.207145) of lysine residues and peptides' N termini, and variable modifications were set to oxidation (+15.994915) of methionine and acetylation (+42.010565) of peptides' N-termini. Precursor ion mass tolerance was set to 10 ppm, and product ion mass tolerance was set to 0.02 Da. The peptide-spectrum-match allowed 1% target false discovery rate (FDR) (strict) and 5% target FDR (relaxed). Files from different batches were analyzed together.

The datasets SP1 and SP3 were produced according to the same protocol as SP2, with following modifications: The TMTpro-labeled peptides were separated into 120 fractions, which were consolidated into 40 fractions. The re-dissolved peptides were analyzed by LC-MS/MS with the same LC system coupled to a Q Exactive HF-X hybrid Quadrupole-Orbitrap (Thermo Fisher Scientific, San Jose, USA) in data dependent acquisition (DDA) mode. The maximum ion injection time for MS1 was set to 50 ms. The mass spectrometric data were analyzed using a protein database composed of the Homo sapiens fasta database downloaded from UniProtKB on 07 Jan 2020, containing 20412 reviewed protein sequences, and the SARS-CoV-2 virus fasta downloaded from NCBI (version NC_045512.2).

The detailed procedures to derive the dataset SP4 was produced according to the same protocol SP2, with following modifications: The TMTpro-labeled peptides were separated into 90 fractions, which were consolidated into 45 fractions. The peptides were dried and redissolved in 2% ACN/0.1% FA, and each of the 45 peptides fractions was injected into the Eksigent NanoLC 400 System (Eksigent, Dublin, CA, USA) coupled with TripleTOF 5600 and 6600 system (SCIEX, CA, USA) for information-dependent acquisition (IDA) MS analysis. All 45 fractions were acquired in a TripleTOF 6600 MS. Then 36 out of 45 fractions were further combined into 12 and analyzed in a TripleTOF 5600 MS. In total, 57 IDA wiff files were analyzed using Maxquant, MSFragger, pFind, and merged to a serum spectral library containing 10001 peptides and 2592 protein groups. We next built a subset library containing 3474 peptides precursors and 536 protein groups for SWATH-MS data analysis using the TripleTOF 5600 system. The SWATH wiff files were converted into mzXML format using msConvert and analyzed using OpenSWATH (version 2.4) against the serum spectral library established above. The retention time extraction window was set to 120 s, and the m/z extraction for MS1 and MS2 was performed at m/z tolerances of 20 and 50 ppm, respectively. The retention time was then calibrated using common internal retention time (CiRT) standards peptides. The m/z extraction for CiRT peptides was performed at an m/z tolerance of 50 ppm. The peptides precursors were identified by OpenSWATH and pyprophet with FDR <0.01. The C-reactive protein (CRP) data were manually analyzed with Skyline. The retention time was predicted by the above-mentioned CiRT peptides, and the isolation time window was set to 2 min. The mass analyzer for MS1 and MS/MS was set to "TOF", with a resolution power of 30000.[12]

The detailed procedures to derive the dataset SP5 follow the same protocol as SP2, with following modifications: After inactivation, ten μL of serum for each sample was depleted using High Select Top-14 Abundant Protein Depletion Mini Spin Columns according to the manufacturer's instructions. The eluates were denatured in 50 μL buffer (8 M urea in 100 mM triethylammonium bicarbonate, TEAB) and were enriched using a 3 kDa Millipore super filtration membrane column. The protein lysates were then reduced, alkylated and digested as above mentioned in SP1 protocol. In each batch, the peptides were fractionated and analyzed as SP1 with minor modifications. In brief, TMT-labeled peptides were separated into 60 fractions and combined into 30 fractions. For MS acquisition, the m/z range of MS1 was 350-1,800 Da with the resolution at 60,000, AGC target was set at 3e6, and maximum ion injection time (max IT) is 50 ms. Top 15 precursors were selected for MS/MS experiment, with the resolution of 45,000, AGC at 2e5, as well as the max IT of 120 ms. The isolation window of selected precursor was 0.7 m/z.[10]

The detailed procedures to generate the proteomic dataset UP1 and UP3 were as follows: urine samples were inactivated and sterilized at 56°C for 30 min and 500 μL urine sample was then precipitated overnight with cold acetone (urine: acetone = 1:4, v/v, −20°C). The precipitated urine samples were centrifuged at 3000 g for 5 min. The protein digestion procedure follows the same protocol as SP2, with following modifications: the digestion was completed by using an enzyme mixture of 5 μg of trypsin and 1 μg of Lys-C at 32°C for 12 h; and the reaction was stopped by adding 110 μL 10% trifluoroacetic acid (TFA).[10]

The dataset UP2 was processed according to the same protocol as UP1 with the following modification: clean peptides were labeled according to the label set (Table S5).

Metabolome analyses

For SM1, ethanol was added to each serum sample (final concentration of 75%) and shaken vigorously to inactivate any potentially existed virus, and the mixture was dried in a biosafety hood. The dried sample was re-dissolved in 80% methanol solution, shaken vigorously and then centrifuged at maximum speed for 10 min. The supernatant containing extracted metabolites was collected, divided into four fractions, and dried under nitrogen gas flow for further metabolomics analysis. The untargeted metabolomics analysis has four different ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) methods, and each method was performed on one of the four fractions of each sample. The UPLC-MS/MS platform consisted of an ACQUITY 2D UPLC system (Waters, Milford, MA, USA) coupled to a QExactive (QE) Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, USA) with HESI-II heated ESI source and Orbitrap mass analyzer. In the first UPLC-MS/MS method, the liquid chromatography column was a C18 reverse phase column (UPLC BEH C18, 2.1 × 100 mm, 1.7 μm; Waters); the mobile solutions were water (A) and methanol (B) containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA); the gradient elution was performed in a 7-min run with B mobile solution gradually increased from 5% to 95%; the QE was operated under positive electron spray ionization mode (+ESI). The second UPLC-MS/MS method was similar to the first method, except that mobile solution B was optimized for more hydrophobic compounds and contained 50% each of methanol and acetonitrile with 0.05% PFPA and 0.01% FA. The third UPLC-MS/MS method also used a C18 column (UPLC BEH C18, 2.1 × 100 mm, 1.7 μm; Waters); the mobile solutions were water (A) and methanol (B) in 6.5 mM ammonium bicarbonate; the QE mass spectrometer was operated under negative ESI mode. The UPLC column used in the fourth method was hydrophilic interaction liquid chromatography (HILIC) column (UPLC BEH Amide, 2.1 × 150 mm, 1.7 μm; Waters), and the mobile solutions were 80% acetonitrile (A) and 10% acetonitrile (B) with 10 mM ammonium formate at pH 10.8; the gradient elution for this method was performed in a 7-min run with A mobile solution decreased from 95% to 5%; the QE was operated under negative ESI mode in the fourth UPLC-MS/MS method. For all four UPLC-MS/MS methods, the mass spectrometer MS scan was operated at a mass resolution of 35,000, the scan range was 70-1,000 m/z, and the analysis was alternated between MS and data-dependent MS2 scan using dynamic exclusion. The MS capillary temperature was 350°C, sheath gas flow rate at 40, aux gas flow rate at 5 for both positive and negative ESI. Raw data pre-processing, peak finding/alignment, and peak annotation were performed using in-house developed software. Metabolites were identified by searching an in-house reference library containing entries of more than 3,000 metabolite reference standards running through the four UPLC-MS/MS methods. Positive identification of a metabolite requires retention index (RI) matching, accurate mass matching with variation less than 10 ppm, and MS/MS spectra matching with high forward and reverse scores of experimental data to library entries. Metabolites identified this way meet the level 1 confidence level by the Chemical Analysis Working Group (CAWG) of the Metabolomics Standards Initiative (MSI).[11]

For UM1, the urine samples were treated the same way as for the serum samples, and each urine sample was also analyzed using the same four UPLC-MS/MS methods used in the metabolomics analysis of the serum.[10,11]

Quality control

We assessed the quality of the MS-based proteomics data via the coefficients of variation for the quantile-normalized values of the protein abundance in pooled samples for the TMTpro-based datasets SP1, SP2, SP5, UP1, and UP2 (Figure S4B), and the PRM-based dataset UP3 (Figure S4C). The Pearson correlation values of technical replicates were used to evaluate the quality of datasets SP3 and SP4 (Figure S4D). For metabolomics analysis, a pooled quality control (PQC) sample was generated by taking an aliquot of each experimental sample and mixing together. This PQC sample was injected multiple times throughout the sample injection sequence. Multiple internal standards (stable isotope labeled metabolites) were spiked into both experimental samples and the PQC samples to monitor platform stability. System variability was determined by calculating the median relative SD (RSD) of the internal standards.[10,11] The missing value ratios distribution for each omics dataset was summarized in Figure S4E.

Machine learning to predict long COVID

The experiment of the multi-omics dataset was developed as an ensemble learning of multiple independent models, corresponding to ten sets of omics data (SM1, SP1, SP2, SP3, SP4, SP5, UM1, UP1, UP2, UP3). The multi-omics and clinical data within 30 days after the disease onset, and 1-year pulmonary and renal long COVID outcome were used for the prediction models. Patients were divided into a multi-omics training set and a test set. Considering that different subsets of patients in the training cohort were enrolled in each omics dataset, we treated each of the ten multi-omics datasets separately during model training. Proteomic results of patients with multiple samplings were averaged. Some samples were randomly selected and analyzed with replicates, showing high reproducibility. Since we averaged the proteomes of multiple samplings, we did not include biological nor technical replicates of multi-omics data in the model. Were replicates included, the weight of these randomly replicated samples would be higher than the rest. Due to the small cohort size, we did not consider sex or age as confounders during training/test set selection. XGBoost method[45] was applied to the training sets in each omics dataset for feature selection and model development. For the 1st step of feature selection, parameters of *subsample, colsample_bytree,* and *colsample_bylevel* were set to 0.5 (as default) to avoid overfitting, scale_pos_weight was set as 10/9 to adjust the imbalance between two classes, and tree number was set as the number of molecules in each omics dataset. Proteins from each dataset were ranked and archived according to their importance. The number of proteins decided the specific number (n) of the selected features with non-zero importance (K1) and the upper limit number (K2) to avoid overfitting (n ≤ K2 and $n \leq 2/3 \times$ K1). For the 2nd step of feature selection, the selected features from each omics dataset were used to construct the corresponding XGBoost classification model with default parameters. Lastly, ten individual models from each

omics dataset were merged into ensemble models. The prediction score of the ensembled model was computed as the average of prediction scores from the inclusive omics dataset classification models.

The experiment for the clinical indicator dataset was developed as follows. Clinical results of patients with multiple samplings were averaged. Apart from the test set, the remaining patients were divided into a clinical indicator training dataset and a validation dataset with a proportion of 2:1. Unlike the feature selection process in omics data, where the upper limit number of features (K2) was empirically determined, we exploited the grid search method for the optimal features and models. The hyperparameter space of the XGBoost algorithm, including feature numbers, learning_rate, subsample, colsample_bytree, colsample_bylevel, and scale_pos_weight, as specified in a way that the optimal ACC and AUC scores on the validation dataset could be searched and the corresponding hyperparameters could be fixed. A classification model was built and achieved using the prediction scores of the test dataset, and the most predictive features were herein obtained.

The predicted score of long COVID ranges from 0 to 1, indicating the probability of development of long COVID. A score higher than 0.5 suggests the patient is more likely to have long COVID.

All the experiments and results were developed with the XGBoost-1.4.2 package on the Python (version: 3.8.10) platform.

Validation analysis

Fifteen patients were recruited from January 26, 2021 to August 30, 2021 for medical examination of chest CT and urine eGFR (Table S6) for serum proteins validation; previously processed peptides of 43 serum samples from 12 patients during 30 days since disease onset were obtained for targeted proteomic analysis. Three pooled samples and three technical replicates were added for quality control. The peptides were spiked with iRT peptides and separated on an 18 cm 75 μm ID fused silica emitter (New Objective) in-house packed with reversed-phase ReproSil-Pur C18-AQ 1.9 μm resin (Dr. Maisch GmbH). The targeted proteins were measured on a nanoflow DIONEX UltiMate 3000 RSLCnano System coupled to a Q Exactive HF hybrid Quadrupole-Orbitrap (both Thermo Fisher Scientific, San Jose, USA) using parallel reaction monitoring (PRM) data acquisition. Peptides were eluted at a flow rate of 300 nL/min with a 30 min gradient (buffer B maintained 10% for 4 min, 10%–30% for 30 min, 30%–80% for 0.1 min, maintained 80% for 2.9 min, 80%–10% for 0.1 min, maintained 10% for 2.9 min). Survey scans were acquired at a resolution of 60,000 with a maximum injection time of 55 ms and a target value of 3e6 over a mass range of 400–2000 m/z. The PRM scans cover the target peptide m/z values from the inclusion list with an isolation window of 1.6 m/z and an acquisition time range of 5 min. The resolution was set to 30,000, the maximum injection time was set to 100 ms, and the target value was set to 2e5. The MS raw data were analyzed by Skyline.[46] Their protein expression was averaged for the patients with multiple samplings during the 30 days. The resultant protein matrix was appended in Table S6.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

The data ranges shown in this paper are represented by the first and third quartiles of the observations unless otherwise mentioned.

For clinical data (demographic information and clinical indicators), p values were calculated by two-sided Wilcoxon rank-sum tests or χ^2 tests unless otherwise mentioned. Different χ^2 tests were chosen based on the total sample size and the theoretical frequency of cells: a) Pearson χ^2 was adopted when the total sample size $n \geq 40$ and the theoretical frequency (expected frequency) $T \geq 5$ of all cells; b) When the total sample size $n \geq 40$ but $1 \leq T < 5$, a continuity correction χ^2 was adopted; c) When the total sample size $n < 40$, or the minimum theoretical frequency $T < 1$, a fisher exact test was adopted. For omics data, p values were calculated by a two-sided Student's t-test. Benjamini & Hochberg correction was applied for p-value adjustment and labeled as adjusted p. Signs: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.005$; ****, $p < 0.001$. R (version 3.6.3) was used for the analysis.

Pathway analysis

String (version: 11.0) was applied to cluster DEPs from renal long COVID. Metascape[47] was used for the functional annotation of serum features in the machine-learning model (Figure 3D). The 894 quantified serum proteins in dataset SP2 were set as the background and all the other parameters were set as default.

ADDITIONAL RESOURCES

This study was registered in the China Clinical Trial Registry with the registration number ChiCTR2100048440: https://www.chictr.org.cn/.