BMJ Open Respiratory Research

Serum KL-6 levels predict clinical outcomes and are associated with *MUC1* polymorphism in Japanese patients with COVID-19

Shuhei Azekawa,[1] Shotaro Chubachi,[1] Takanori Asakura,[1,2,3] Ho Namkoong,[4] Yasunori Sato,[5] Ryuya Edahiro,[6,7] Ho Lee,[1] Hiromu Tanaka,[1] Shiro Otake,[1] Kensuke Nakagawara,[1] Takahiro Fukushima,[1] Mayuko Watase,[1] Kaori Sakurai,[1] Tatsuya Kusumoto,[1] Katsunori Masaki ,[1] Hirofumi Kamata,[1] Makoto Ishii,[8] Naoki Hasegawa,[4] Yukinori Okada,[6,9,10,11,12,13] Ryuji Koike,[14] Yuko Kitagawa,[15] Akinori Kimura,[16] Seiya Imoto,[17] Satoru Miyano,[18] Seishi Ogawa,[19] Takanori Kanai,[20] Koichi Fukunaga,[1] The Japan COVID-19 Task Force

To cite: Azekawa S, Chubachi S, Asakura T, *et al*. Serum KL-6 levels predict clinical outcomes and are associated with *MUC1* polymorphism in Japanese patients with COVID-19. *BMJ Open Resp Res* 2023;**10**:e001625. doi:10.1136/bmjresp-2023-001625

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjresp-2023-001625).

SC and TA contributed equally.

Received 11 January 2023
Accepted 17 May 2023

For numbered affiliations see end of article.

Correspondence to
Dr Shotaro Chubachi;
bachibachi472000@z6.keio.jp

ABSTRACT

Background Krebs von den Lungen-6 (KL-6) is a known biomarker for diagnosis and monitoring of interstitial lung diseases. However, the role of serum KL-6 and the mucin 1 (*MUC1*) variant (rs4072037) in COVID-19 outcomes remains to be elucidated. We aimed to evaluate the relationships among serum KL-6 levels, critical outcomes and the *MUC1* variant in Japanese patients with COVID-19.

Methods This is a secondary analysis of a multicentre retrospective study using data from the Japan COVID-19 Task Force collected from February 2020 to November 2021, including 2226 patients with COVID-19 whose serum KL-6 levels were measured. An optimal serum KL-6 level cut-off to predict critical outcomes was determined and used for multivariable logistic regression analysis. Furthermore, the relationship among the allele dosage of the *MUC1* variant, calculated from single nucleotide polymorphism typing data of genome-wide association studies using the imputation method, serum KL-6 levels and COVID-19 critical outcomes was evaluated.

Results Serum KL-6 levels were significantly higher in patients with COVID-19 with critical outcomes (511±442 U/mL) than those without (279±204 U/mL) (p<0.001). Serum KL-6 levels ≥304 U/mL independently predicted critical outcomes (adjusted OR (aOR) 3.47, 95% CI 2.44 to 4.95). Moreover, multivariable logistic regression analysis with age and sex indicated that the *MUC1* variant was independently associated with increased serum KL-6 levels (aOR 0.24, 95% CI 0.28 to 0.32) but not significantly associated with critical outcomes (aOR 1.11, 95% CI 0.80 to 1.54).

Conclusion Serum KL-6 levels predicted critical outcomes in Japanese patients with COVID-19 and were associated with the *MUC1* variant. Therefore, serum KL-6 level is a potentially useful biomarker of critical COVID-19 outcomes.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Krebs von den Lungen-6 (KL-6) is a candidate biomarker for diagnosis and monitoring of interstitial lung diseases.

WHAT THIS STUDY ADDS

⇒ High serum KL-6 levels (≥304 U/mL) were significantly associated with and independently predicted critical outcomes in patients with COVID-19.

⇒ The mucin 1 (*MUC1*) variant was found to be independently associated with increased serum KL-6 levels, but was not significantly associated with critical outcomes.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our study provides substantial evidence that serum KL-6 levels and the *MUC1* variant may be useful biomarkers of critical COVID-19 outcomes.

⇒ This could improve treatment outcomes of patients with severe COVID-19.

INTRODUCTION

COVID-19 was first discovered in China in December 2019 and has since spread globally,[1] infecting more than 600 million people and killing 6.5 million.[2] Despite great advances in the development of vaccines and treatments, the disease continues to affect numerous individuals worldwide. Therefore, developing biomarkers that can predict COVID-19 severity is essential for effectively using healthcare resources and identifying targets for new therapeutic agents.[3]

Krebs von den Lungen-6 (KL-6) is a mucin-like, high-molecular-weight glycoprotein classified as a mucin 1 (*MUC1*) antigen.[4 5] Serum KL-6 levels are a biomarker for diagnosis and monitoring of disease activity in patients with various interstitial lung diseases (ILDs).[6] Its

levels are also increased in various respiratory infections, including pneumocystis pneumonia,[7] mycobacterial infections[8 9] and viral infection.[10] Since the beginning of the COVID-19 pandemic, elevated serum KL-6 levels have been reported to be associated with COVID-19 severity.[11–15] Moreover, the extent of lung lesions in COVID-19-associated pneumonia is related to increased serum KL-6 levels and poor prognosis.[16 17] However, some studies have found no association between KL-6 levels and disease severity.[18–20] Therefore, the clinical role of KL-6 in COVID-19 is still controversial, which might be a consequence of the varied and limited sample sizes of the aforementioned studies.

KL-6 levels vary following *MUC1* gene polymorphisms and demonstrate racial differences.[21 22] In ILDs, *MUC1* polymorphisms have been reported to be associated with disease susceptibility.[21–25] However, no studies conducted so far have focused on the relationships among serum KL-6 levels, KL-6-associated genetic polymorphisms and COVID-19 severity. The Japan COVID-19 Task Force was founded in early 2020 as a national multicentre consortium to combat COVID-19. Since February 2020, more than 100 institutions across Japan have contributed clinical data and specimens from patients with COVID-19 to improve knowledge of this disease.[26–31] Using this large pool of available data, the present study aimed to (1) determine whether serum KL-6 is associated with important clinical outcomes, such as mortality, extracorporeal membrane oxygenation (ECMO) and invasive positive pressure ventilation (IPPV), that more directly reflect COVID-19 prognosis than its severity; and (2) clarify the association among serum KL-6 levels, important critical outcomes and KL-6-associated genetic polymorphisms.

METHODS

Study design and setting

Data for all COVID-19 cases used for the present secondary analysis of a multicentre retrospective study were obtained by the Japan COVID-19 Task Force from February 2020 to September 2021.[26] Patients were at least 18 years old and diagnosed with COVID-19 based on detection of SARS-CoV-2 via PCR or antigen analysis. The data of patients who agreed to participate in the present study were registered in an electronic case record form. All patients who participated in the present study provided written or oral informed consent.

Figure 1 shows the enrolment process. Of the 3424 patients enrolled, 1198 were excluded because they lacked clinical information on KL-6 levels (956 patients), were non-Japanese (69 patients), presented malignancy (165 patients) or presented an ILD as a comorbidity (8 patients).[6 32] Finally, 2226 patients were analysed. For assessing KL-6-associated genetic polymorphisms, 1119 patients available for genome-wide association study (GWAS) were included.

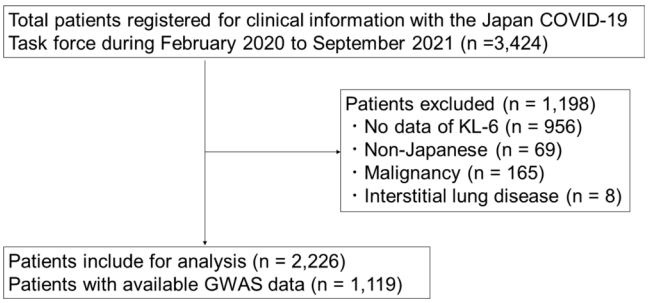

Figure 1 Population flow chart of the study cohort. 3424 patients with COVID-19 were hospitalised during the study period. 1198 patients were excluded for reasons indicated in the figure. Therefore, 2226 patients were included for analysis and 1119 patients for whom GWAS data were available were included in the SNP analysis. GWAS, genome-wide association studies; KL-6, Krebs von den Lungen-6; SNP, single nucleotide polymorphism.

Data collection and definition

The following clinical information was extracted from the electronic case record forms: age, gender, height, weight, smoking history, comorbidities, symptoms and signs at initial presentation, blood test results, and radiological test results. Symptoms and signs included those detected at the time of referral and on admission and those detected during hospitalisation. Blood results and radiographic imaging using chest X-ray (CXR) and CT were collected within 48 hours of admission. The attending physician at each facility determined when to perform the follow-up CXR. Rapid deterioration of CXR was defined as a deterioration of lung infiltrates in more than 50% of the lung fields within 48 hours compared with the CXR on admission, based on the criteria for severe disease in COVID-19-positive patients.[33] Serum KL-6 levels were measured using the Nanopia KL-6 Reagent Kit (Sekisui Medical, Tokyo, Japan). KL-6 levels were measured in the serum within 48 hours of admission, but longitudinal KL-6 levels were not collected. COVID-19 severity was defined as follows: most severe, patients requiring support using high-flow oxygen devices, IPPV or ECMO, or cases leading to death; severe, patients requiring support using low-flow oxygen devices; and asymptomatic/mild, asymptomatic/symptomatic patients not requiring oxygen support.[26] To investigate the association between serum KL-6 levels and clinically important outcomes, critical outcomes were defined as follows: use of IPPV, ECMO or high-flow oxygen supply or patient death during hospitalisation.[27]

Genotype imputation

GWAS genotyping was performed on samples from 2520 COVID-19 cases using the Infinium Asian Screening Array (Illumina, San Diego, California, USA). We applied stringent quality control (QC) filters to the samples and variants. The 2393 COVID-19 cases passed the sample QC. Details of the QC are described elsewhere.[26] After QC, we performed genome-wide genotype imputation. Details of

Table 1 Baseline characteristics of patients included for analysis

Variable	Total (n =2226)	Critical outcome (–) (n=1792)	Critical outcome (+) (n=434)	P value
Age, years	56.3±17.0	54.5±17.4	63.5±12.8	<0.001
Gender, male	1529 (68.7)	1205 (67.2)	324 (74.7)	0.003
BMI, kg/m^2	24.8±4.8	24.6±4.6	25.8±5.4	<0.001
Current smoker	328 (15.7)	268 (15.85)	60 (15.0)	0.703
Smoking history	982 (47.2)	772 (45.8)	210 (53.3)	0.024
Onset to hospitalisation, days	7.03±4.50	6.53±4.15	8.77±5.18	<0.001
Comorbidities				
Hypertension	741 (33.7)	520 (29.0)	221 (51.5)	<0.001
Diabetes mellitus	467 (21.2)	313 (17.5)	154 (35.7)	<0.001
Cardiovascular disease	214 (9.7)	150 (8.4)	64 (14.7)	<0.001
Autoimmune disease	90 (4.1)	78 (4.4)	12 (2.8)	0.125
COPD	85 (3.9)	54 (3.1)	31 (7.2)	<0.001
Asthma	150 (6.9)	124 (7.1)	26 (6.1)	0.459
Hyperuricaemia	231 (10.5)	162 (9.2)	69 (16.0)	<0.001
Chronic liver disease	99 (4.6)	77 (4.5)	22 (5.2)	0.535
Chronic kidney disease	159 (7.5)	91 (5.0)	68 (15.7)	<0.001
Treatment				
Antibiotics	568 (25.7)	289 (16.2)	279 (65.2)	<0.001
Favipiravir	575 (26.0)	439 (24.7)	136 (31.7)	0.003
Remdesivir	863 (39.2)	564 (31.8)	299 (69.5)	<0.001
Tocilizumab	269 (12.2)	105 (5.9)	164 (38.6)	<0.001
Anticoagulant	693 (31.4)	326 (18.3)	367 (85.8)	<0.001
Systemic corticosteroid	1168 (52.9)	764 (42.9)	404 (94.2)	<0.001
Critical outcome	434 (19.5)	0 (0)	434 (100.0)	<0.001
High-flow oxygen therapy	104 (4.7)	0 (0)	104 (24.0)	<0.001
IPPV	52 (2.4)	0 (0)	52 (12.2)	<0.001
ECMO	317 (14.3)	0 (0)	317 (73.0)	<0.001
Mortality	77 (3.5)	0 (0)	77 (17.7)	<0.001

Data are presented as mean±SD or n (%) of patients. Differences in variables with and without critical events were compared using unpaired t-test or χ^2 test.
BMI, body mass index; COPD, chronic obstructive pulmonary disease; ECMO, extracorporeal membrane oxygenation; IPPV, invasive positive pressure ventilation.

genotype imputation are described elsewhere.[26] Out of 2393 patients with COVID-19, 1119 cases had information on rs4072037 mutation. The mutation of rs4072037 in *MUC1* is known to be associated with serum KL-6 levels.[21 22] The dosage effects of rs4072037 on serum KL-6 levels were evaluated using a linear regression model with age and sex as covariates.

Statistical analysis

For baseline characteristics, summary statistics were performed using frequencies and proportions for categorical data, and means and SD for continuous variables. For comparisons between the groups with high and low KL-6 levels, unpaired t-test or Wilcoxon rank-sum test was used for continuous variables, while χ^2 test or Fisher's exact test was used for categorical variables. The Dunnett's multiple comparisons test was used to compare serum KL-6 levels among the three groups classified following the type of chest imaging findings. Receiver operating characteristic (ROC) curve analysis was performed to determine appropriate serum KL-6 cut-off values for critical outcomes using the Youden index. To investigate the association between KL-6 levels and critical outcomes, a multivariable logistic regression analysis was performed, adjusting for the following clinical characteristics: age, sex, hypertension, haemoglobin A1c (HbA1c) ≥6.5%, cardiovascular disease, chronic obstructive pulmonary disease (COPD), asthma, chronic

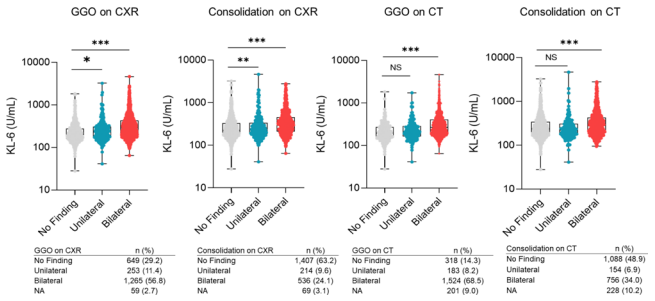

Figure 2 Differences in serum KL-6 levels based on the chest imaging findings on admission. Group comparisons were performed using Dunnett's multiple comparisons test. *P<0.05, **P<0.01, ***P<0.001. CXR, chest X-ray; GGO, ground glass opacity; KL-6, Krebs von den Lungen-6.

liver disease and chronic kidney disease.[34–36] To elucidate the clinical relevance of serum KL-6 levels, we calculated the Pearson correlation coefficients between KL-6 and inflammatory biomarkers and performed multivariable logistic regression analyses for critical outcomes using these inflammatory biomarkers as covariates, in addition to serum KL-6 and previously reported risk factors.[34–36] To evaluate the association between rs4072037 genotypes and serum KL-6 levels, a multivariable logistic regression analysis was conducted with age and sex. P values of 0.05 or less were considered statistically significant. All statistical analyses were performed using SPSS Statistics V.28 and GraphPad Prism V.9.0 (GraphPad Prism, La Jolla, California, USA).

RESULTS

Baseline characteristics of patients

Patients with COVID-19 whose serum KL-6 levels were measured on admission had more severe disease and more frequently had critical outcomes than those whose serum KL-6 levels were not measured (online supplemental table 1). The baseline clinical characteristics of the 2226 patients included in the analysis are shown in table 1. The mean age of the patients was 56.3 years (SD, 17.0), and 1529 were male (68.7%). The 434 patients (19.5%) experiencing a critical outcome during hospitalisation were significantly older than the patients without a critical outcome (63.5 years vs 54.5 years) and were more frequently male than female (74.7% vs 67.2%). The group experiencing critical outcomes was also significantly more likely to show comorbidities, such as hypertension (50.9% vs 29.0%), diabetes mellitus (33.0% vs 14.1%), cardiovascular disease (14.7% vs 8.4%), COPD (7.1% vs 3.0%), hyperuricaemia (15.9% vs 9.0%) and chronic kidney diseases (15.7% vs 5.0%), than the group not experiencing such outcomes. Figure 2 shows the differences in serum KL-6 levels based on chest imaging performed within 48 hours of admission. The group with bilateral ground glass opacity (GGO) and consolidation on CXR had significantly higher serum KL-6 levels than the group without such characteristics. Regarding CT images, patients with bilateral but not unilateral GGO or CXR showed a significant increase in KL-6 levels compared with patients not showing bilateral GGO.

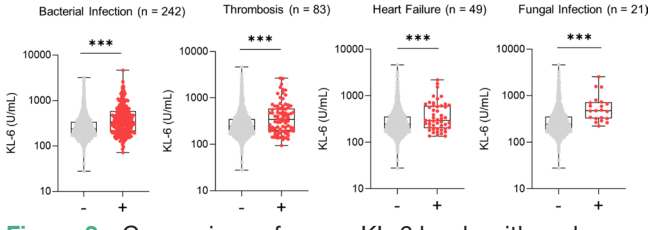

Figure 3 Comparison of serum KL-6 levels with and without post-hospitalisation complications (bacterial infection, thrombosis, heart failure and fungal infection). Unpaired t-test was performed. ***P<0.001. KL-6, Krebs von den Lungen-6.

KL-6 levels predict in-hospital complications

Figure 3 depicts the serum KL-6 levels following the presence or absence of complications during hospitalisation. Serum KL-6 levels were significantly higher in the group with bacterial and/or fungal infection, thrombosis and heart failure than in the group without these complications.

KL-6 levels predict critical outcomes

Regarding disease severity on admission, patients with severe or most severe disease had significantly higher serum KL-6 levels than those with asymptomatic/mild disease (online supplemental figure 1). Serum KL-6 levels were also higher in patients with a rapid deterioration of CXR findings within 48 hours of admission compared with those without (online supplemental figure 1). Figure 4A depicts the serum KL-6 levels following COVID-19 severity. Serum KL-6 levels were significantly higher in patients with severe and most severe COVID-19 than in those with mild COVID-19 severity (most severe 510±443 U/mL vs mild 248±162 U/mL; severe 333±163 U/mL vs mild 248±162 U/mL). Serum KL-6 levels were significantly higher in patients with critical outcomes than in those without them (figure 4B). Figure 4C shows the oxygen demand and mortality on admission and at the time of the most severe illness for patients with high and low serum KL-6 on admission. It was shown that nearly half of the patients with elevated serum KL-6 levels on admission have a higher oxygen demand than on admission and that the oxygen demand tended to increase over time. In addition, ROC curve analysis was performed to evaluate the diagnostic value of KL-6 levels for the critical outcomes. The optimal cut-off value of KL-6 levels for the critical outcomes was 304 U/mL, with an area under the curve of 0.73 (95% CI 0.71 to 0.77) (figure 4D). Each critical outcome was consistent with this result (online supplemental figure 2). Even after excluding patients with critical outcomes on admission, serum KL-6 levels were associated with subsequent poor outcomes (online supplemental table 2). Clinical characteristics were

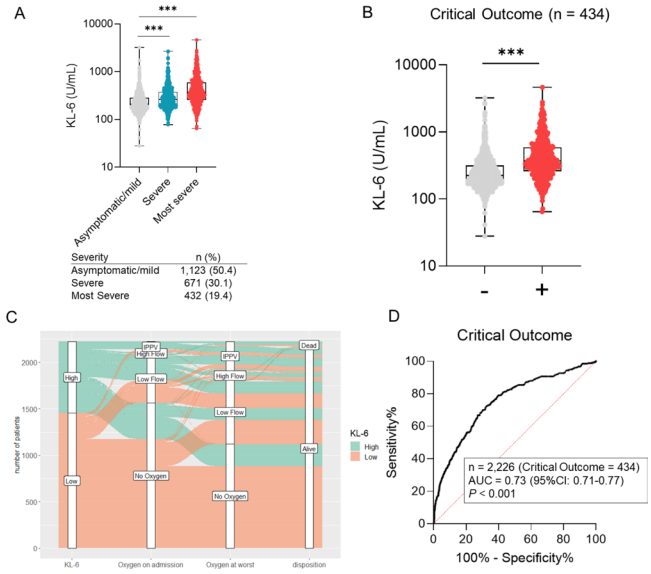

Figure 4 (A) Comparison of serum KL-6 levels following the severity of COVID-19. Group comparisons were performed using Dunnett's multiple comparisons tests. (B) Serum KL-6 levels of patients with or without critical outcomes. Unpaired t-tests were performed. (C) Alluvial plot showing trends on admission, worst oxygen demand and mortality by serum KL-6 level on admission. 'KL-6 High' means serum KL-6 ≥304 U/mL and 'KL-6 Low' means serum KL-6 <304 U/mL. 'High Flow' means that high-flow nasal cannula or non-invasive positive pressure ventilation was used for treatment. 'Low Flow' means some oxygen demand without IPPV or high-flow nasal cannula or non-invasive positive pressure high-flow nasal cannula or non-invasive positive pressure was used for treatment. (D) Receiver operating characteristic curve for outcomes of COVID-19 based on the serum levels of KL-6. ***P<0.001. AUC, area under the curve; IPPV, invasive positive pressure ventilation; KL-6, Krebs von den Lungen-6.

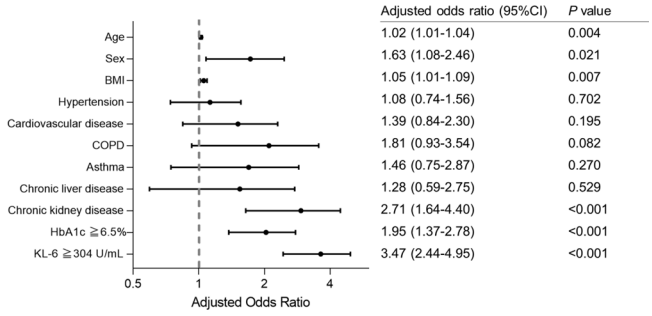

Figure 5 Multivariable logistic regression analysis for ascertaining the relationship between critical outcomes and serum KL-6 and already known risk factors for COVID-19. Forest plot showing the adjusted OR, 95% CI and p value of the parameters of the analysis. BMI, body mass index; COPD, chronic obstructive pulmonary disease; HbA1c, haemoglobin A1c; KL-6, Krebs von den Lungen-6.

compared by dividing the patients into two groups: the high KL-6 group (KL-6 ≥304 U/mL) and the low KL-6 group (KL-6 <304 U/mL) (online supplemental table 3). Concerning clinical signs, the high KL-6 group had significantly higher rates of unconsciousness and shortness of breath, both suggesting severe COVID-19, than the low KL-6 group. Regarding laboratory data, white cell count, lactate dehydrogenase, uric acid, ferritin, D-dimer and C reactive protein (CRP) levels were significantly higher in the high KL-6 group than in the low KL-6 group. These serological biomarkers are known to reflect inflammation in patients with COVID-19.[37–39]

The multivariable logistic regression analysis revealed that high KL-6 levels (adjusted OR (aOR) 3.47, 95% CI 2.44 to 4.95) as well as older age (aOR 1.02, 95% CI 1.01 to 1.04), male status (aOR 1.63, 95% CI 1.08 to 2.46), increased body mass index (aOR 1.05, 95% CI 1.01 to 1.09), chronic kidney disease (aOR 2.71, 95% CI 1.64 to 4.40) and HbA1c levels ≥6.5% (aOR 1.95, 95% CI 2.44 to 4.95) were independently associated with critical outcomes (figure 5). Using the same covariates as for critical outcomes, multivariable logistic regression analyses showed that KL-6 levels are independently associated with IPPV (aOR 3.81, 95% CI 2.59 to 5.61) and ECMO (aOR 6.41, 95% CI 2.56 to 16.05), but not mortality (aOR 1.52, 95% CI 0.77 to 3.03) (online supplemental figure 3). Serum CRP, ferritin and D-dimer levels were weakly correlated with serum KL-6 levels (online supplemental table 4). Multivariable logistic regression analysis also showed that KL-6 was independently associated with CRP, ferritin and D-dimer for critical outcomes (online supplemental figure 4).

Clinical relevance of *MUC1* polymorphisms

Table 2 shows that the allele dosage of the *MUC1* variant (rs4072037) adjusted for age and sex was significantly associated with serum KL-6 levels (aOR 0.24, 95% CI 0.28 to 0.32). This suggested that the presence of the C allele, a reference allele of rs4072037, was associated with increased KL-6 levels.

DISCUSSION

The present study provided two novel findings with clinical relevance. First, serum KL-6 levels could predict the clinical outcomes of Japanese patients with COVID-19 in this large multicentre study; additionally, we successfully determined the optimal cut-off value of serum KL-6 levels to predict such critical outcomes. It has already been reported that serum KL-6 effectively predicts the severity of COVID-19 or fibrotic change in patients with severe COVID-19.[40 41] However, in a large patient population including patients with mild disease, for the first time in this study, KL-6 was stratified according to disease severity and was shown to be associated with hard endpoints. The present study also allowed us to perform a multivariate analysis where we showed an association between KL-6 levels and clinical outcomes using confounding factors, including biomarkers other than serum KL-6 and known risk factors.[34–39] Moreover, the results showed that serum KL-6 levels were related to complications during hospitalisation. Therefore, the serum KL-6 levels of patients

Table 2 Multivariable logistic regression analysis of allele dosage of rs4072037 and serum KL-6 levels

	KL-6			Critical outcomes		
Variable	aOR	95% CI	P value	aOR	95% CI	P value
Allele dosage	0.24	0.18 to 0.32	<0.001	1.11	0.80 to 1.54	0.547
Age	1.05	1.04 to 1.06	<0.001	1.06	1.04 to 1.07	<0.001
Sex, male	1.78	1.29 to 2.64	<0.001	2.1	1.41 to 3.14	<0.001

Adjusted OR was estimated using logistic regression with adjustments for age and sex.
aOR, adjusted OR; KL-6, Krebs von den Lungen-6.

with COVID-19 may be a clinical indicator for providing more aggressive treatment. Second, the *MUC1* variant (rs4072037) was related to serum KL-6 levels but not to the critical outcomes of patients with COVID-19. To the best of our knowledge, this is the first study to examine the association of *MUC1* polymorphism with serum KL-6 levels and COVID-19 clinical outcomes.

The present study provides strong evidence that KL-6 levels can predict the critical outcomes of COVID-19 through significant multicentre validation. However, the efficacy of several serological biomarkers, such as those related to the inflammatory response and coagulation predisposition, in predicting COVID-19 critical outcomes has been previously reported.[37–39] Notably, our results also showed that KL-6 was independently associated with critical outcomes, even when adding CRP, ferritin and D-dimer, which are known serological biomarkers of COVID-19, to the multivariable analysis. This suggests that serum KL-6 levels can be a complementary measurement of these biomarkers (online supplemental figure 4).[37 38] Serum KL-6 is a known biomarker of lung injury and fibrosis, reflecting damage or regeneration of type 2 pneumocytes,[42] based on its previously reported association with the severity and prognosis of ILDs and acute respiratory distress syndrome (ARDS).[43–45] In the present study, KL-6 levels were also associated with the degree of COVID-19-derived pneumonia on imaging evaluation. A possible mechanism underlying the increased serum KL-6 levels in severe COVID-19 cases is the increased production of type 2 pneumocytes due to direct viral infection and leakage into the blood due to disruption of the alveolar–capillary barrier.[40] Patients with severe COVID-19 often end up having ARDS, and deaths due to COVID-19 are associated with a higher incidence of ARDS than that in COVID-19 survivors. It has been reported that autopsy samples of the lungs of patients with COVID-19 show a pattern of diffuse alveolar damage (DAD), a pathology of ARDS.[46 47] Therefore, it can be inferred that the primary mechanism underlying the increase in serum KL-6 levels is the leakage of KL-6 from the lungs into the blood due to virus proliferation and destruction of the alveolar epithelium and basement membrane with DAD.[40] In fact, not only blood but also epithelial lining fluid KL-6 levels are increased in patients with ARDS,[48] and high levels of this biomarker have also been reported in the bronchoalveolar lavage fluid of patients with severe COVID-19.[49] In the present study, KL-6 levels were associated with bilateral shadows on imaging findings, consistent with the results of previous studies.[16 50]

The KL-6 cut-off value for predicting critical outcomes was 304 U/mL. This value was lower than the cut-off values found for diagnosing ILDs and poor prognostic factors for ARDS and ILDs.[45 51] Furthermore, the KL-6 cut-off value was lower than that previously reported for predicting the severity of COVID-19.[12 14 41] We hypothesise that this may be because previous studies had fewer participants but included more patients with relatively severe disease compared with that in the present study. The alveolar–capillary permeability in COVID-19 is elevated in the acute phase and may further increase with lung fibrosis.[12] A previous report on COVID-19 suggested that peak serum KL-6 levels during hospitalisation are more helpful than those at diagnosis in predicting clinical outcomes.[12] Therefore, further evaluation of serum KL-6 levels over time should be explored in future studies.

The present study also revealed the relationships between the *MUC1* variant and serum KL-6 levels in Japanese patients with COVID-19. The results obtained here are consistent with those of previous reports demonstrating that the *MUC1* variant affects KL-6 levels in ILDs and sarcoidosis.[21 22] MUC1 is an extracellular protein anchored to the epithelial surface and involved in morphogenetic signalling.[21 22 52] The presence of the rs4072037 SNP (single nucleotide polymorphism) causes alternative splicing of the exon regions under its regulation, leading to abnormal transcription.[53] MUC1 has been demonstrated to be a critical innate immunity modulator,[54] acting as an important and necessary anti-inflammatory agent during airway infection.[25 55] It is controversial whether the *MUC1* variant affects the severity of ILDs. In support of this hypothesis, it has been reported that the *MUC1* variant affects the severity of pulmonary alveolar proteinosis.[24] However, some studies showed that rs4072037 may be involved in disease susceptibility to ILDs but is not related to the severity of the disease.[23] For example, high frequency of the rs4072037 C allele and serum KL-6 levels have been reported in patients with antisynthetase syndrome compared with those in healthy control subjects.[23] It has also been reported that in patients with systemic sclerosis-ILD, the rs4072037 mutation is significantly associated with increased serum KL-6 levels but is not a predictor of a

lower diffusing capacity for carbon monoxide.[23] In the present study, the *MUC1* variant did not correlate with COVID-19 severity. Moreover, it has been reported that the optimal cut-off value of serum KL-6 levels discriminating patients with ILD from healthy control subjects differs following the rs4072037 genotype.[21] The results of the present study suggest that high serum KL-6 levels may be due not only to severe COVID-19, but also to the *MUC1* variant, which may result in false-positive serum KL-6 in patients with severe COVID-19. Although it is not possible to measure *MUC1* variants in actual clinical practice at this time, caution should be exercised in interpreting serum KL-6 in patients with COVID-19.

Recently, it has been recognised that COVID-19 symptoms can persist after the disease, which is referred to as long COVID or postacute COVID-19 syndrome.[56 57] It is estimated that 22.9% of patients with COVID-19 have residual shortness of breath.[58] It has also been reported that within 1–6 months of discharge from the hospital, 55.7% of patients have residual abnormal findings on CT scans and 44.3% have residual abnormal pulmonary function tests.[59] Furthermore, 56% of patients treated with mechanical ventilation have been reported to have reduced pulmonary diffusion capacity 6 months after discharge.[60] Serum KL-6 levels may also help predict and diagnose long COVID. An association between residual CT shadows and KL-6 levels at 3 months has been reported.[17] Future studies should examine the relationship among serum KL-6 levels, changes on imaging and long COVID on admission, during hospitalisation and after discharge.

There are several limitations to the present study. First, image scoring was not available. Previous studies have reported a correlation between CT quantitative pneumonia range and serum KL-6 levels.[61 62] In this large, multicentre study, CT could only be evaluated qualitatively; future comparisons with quantitative evaluations are desirable. Second, serum KL-6 levels were not measured over time. Serum KL-6 levels continue to rise during hospitalisation and peak later in severe cases than in mild cases[62 63]; moreover, peak serum KL-6 levels have been reported to be more useful in predicting COVID-19 severity and death than serum KL-6 levels on admission.[12 41 64] Although measuring peak levels may improve sensitivity and specificity, the fact that serum KL-6 levels on admission were predictive of outcome is both convenient and clinically useful. Third, critical outcomes are significantly higher in patients whose serum KL-6 levels were measured on admission than in those who did not. It is speculated that this result might be because KL-6 was aggressively measured in patients with more severe COVID-19 on admission to assess the extent of lung disease. This selection bias may affect the prediction values of serum KL-6 values for critical outcomes. Fourth, in this study, we were unable to show the prevalence and severity of ARDS with serum KL-6 levels because we could not assess PaO_2 (partial pressure of arterial oxygen) levels via arterial blood gas analysis on admission. Severe COVID-19 is associated with ARDS, but the effect of ARDS on serum KL-6 levels of patients with COVID-19 cannot be determined from this study. However, we revealed that serum KL-6 levels were higher in patients with more severe disease on admission or those with rapid deterioration. These findings support that serum KL-6 is a predictive marker in patients with ARDS-like severe COVID-19 or rapidly deteriorating COVID-19 on early admission.

In conclusion, the levels of serum KL-6 within 48 hours of admission were associated with critical outcomes in a large Japanese multicentre study involving patients with COVID-19. Further, the *MUC1* variant was associated with serum KL-6 levels within 48 hours of admission but not with critical outcomes. Thus, serum KL-6 levels on admission may be a useful biomarker of critical COVID-19 outcomes.

Author affiliations
[1]Division of Pulmonary Medicine, Department of Medicine, Keio University School of Medicine Graduate School of Medicine, Tokyo, Japan
[2]Department of Clinical Medicine (Laboratory of Bioregulatory Medicine), Kitasato University School of Pharmacy, Tokyo, Japan
[3]Department of Respiratory Medicine, Kitasato University, Kitasato Institute Hospital, Tokyo, Japan
[4]Department of Infectious Diseases, Keio University School of Medicine, Tokyo, Japan
[5]Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
[6]Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita, Japan
[7]Department of Respiratory Medicine and Clinical Immunology, Osaka University Graduate School of Medicine, Suita, Japan
[8]Department of Respiratory Medicine, Nagoya University Graduate School of Medicine Faculty of Medicine, Nagoya, Japan
[9]Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Suita, Japan
[10]Center for Infectious Disease Education and Research (CiDER), Osaka University, Suita, Japan
[11]Laboratory for Systems Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan
[12]Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita, Japan
[13]Department of Genome Informatics, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan
[14]Medical Innovation Promotion Center, Tokyo Medical and Dental University, Tokyo, Japan
[15]Department of Surgery, Keio University School of Medicine, Tokyo, Japan
[16]Institute of Research, Tokyo Medical and Dental University, Tokyo, Japan
[17]Division of Health Medical Intelligence, Human Genome Center, the Institute of Medical Science, The University of Tokyo, Tokyo, Japan
[18]M&D Data Science Center, Tokyo Medical and Dental University, Tokyo, Japan
[19]Department of Pathology and Tumor Biology, Kyoto University Graduate School of Medicine Faculty of Medicine, Kyoto, Japan
[20]Division of Gastroenterology and Hepatology, Department of Internal Medicine, Keio University School of Medicine Graduate School of Medicine, Tokyo, Japan

Acknowledgements We would like to thank all participants involved in this study and all members of the Japan COVID-19 Task Force who regularly engaged in clinical and research work on COVID-19.

Collaborators The list of all members who contributed to this study is shown in the online supplemental information.

Contributors SA, SC, TA, HN, KM, HK, MI and KF conceptualised the present study. SA, SC, TA, HN, HL, HT, SOt, KN, TF, MW, KS and TKu performed data curation and interpretation. SA, SC and TA performed data analyses and visualisation. YS supervised the statistical analysis. YO and RE performed the genotype imputation. SA, SC and TA drafted the paper. HN, YS, RE, HL, HT, SOt, KN, TF, MW, KS, TKu, KM, HK, MI, NH, YO, RK, YK, AK, SI, SM, SOg, TKa and KF reviewed and edited the paper. All authors approved the final draft of the manuscript for publication. SC accepts full responsibility for the work and the conduct of the study, accesses to the data, and controls the decision to publish.

Funding This study was supported by the AMED (JP20nk0101612, JP20fk0108415, JP21jk0210034, JP21km0405211, JP21km0405217, JP21wm0325031 and JP21fk0108573), JST CREST (JPMJCR20H2), JST PRESTO (JPMJPR21R7) and MHLW (20CA2054).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval This study involves human participants and was approved by the Ethics Committee of Keio University School of Medicine and related research institutions (#20200061). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available.

ORCID iD

Katsunori Masaki http://orcid.org/0000-0003-0909-9409

REFERENCES

1 Xiong J, Lipsitz O, Nasri F, *et al*. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. *J Affect Disord* 2020;277:55–64.
2 WHO. WHO Coronavirus (COVID-19) dashboard. 2020. Available: https://covid19.who.int/ [Accessed 11 Dec 2022].
3 Ponti G, Maccaferri M, Ruini C, *et al*. Biomarkers associated with COVID-19 disease progression. *Crit Rev Clin Lab Sci* 2020;57:389–99.
4 Kohno N, Kyoizumi S, Awaya Y, *et al*. New serum indicator of interstitial pneumonitis activity. Sialylated carbohydrate antigen KL-6. *Chest* 1989;96:68–73.
5 Kohno N, Awaya Y, Oyama T, *et al*. KL-6, a Mucin-like glycoprotein, in bronchoalveolar lavage fluid from patients with interstitial lung disease. *Am Rev Respir Dis* 1993;148:637–42.
6 Ishikawa N, Hattori N, Yokoyama A, *et al*. Utility of KL-6/MUC1 in the clinical management of interstitial lung diseases. *Respir Investig* 2012;50:3–13.
7 Urabe N, Sakamoto S, Sano G, *et al*. Serial change in serum biomarkers during treatment of non-HIV pneumocystis pneumonia. *J Infect Chemother* 2019;25:936–42.
8 Inoue Y, Nishimura K, Shiode M, *et al*. Evaluation of serum KL-6 levels in patients with pulmonary tuberculosis. *Tuber Lung Dis* 1995;76:230–3.
9 Asakura T, Kimizuka Y, Nishimura T, *et al*. Serum Krebs von den Lungen-6 level in the disease progression and treatment of Mycobacterium Avium complex lung disease. *Respirology* 2021;26:112–9.
10 Kawasaki Y, Aoyagi Y, Abe Y, *et al*. Serum KL-6 levels as a biomarker of lung injury in respiratory syncytial virus bronchiolitis. *J Med Virol* 2009;81:2104–8.
11 Naderi N, Rahimzadeh M. Krebs von den Lungen-6 (KL-6) as a clinical marker for severe COVID-19: a systematic review and meta-analyses. *Virology* 2022;566:106–13.
12 Awano N, Inomata M, Kuse N, *et al*. Serum KL-6 level is a useful biomarker for evaluating the severity of Coronavirus disease 2019. *Respir Investig* 2020;58:440–7.
13 d'Alessandro M, Bergantini L, Cameli P, *et al*. Peripheral biomarkers' panel for severe COVID-19 patients. *J Med Virol* 2021;93:1230–2.
14 d'Alessandro M, Cameli P, Refini RM, *et al*. Serum KL-6 concentrations as a novel biomarker of severe COVID-19. *J Med Virol* 2020;92:2216–20.
15 Japan ECM. Onet for COVID-19. Nationwide system to centralize decisions around ECMO use for severe COVID-19 pneumonia in Japan (special correspondence). *J Intensive Care* 2020;8:29.
16 Xue M, Zheng P, Bian X, *et al*. Exploration and correlation analysis of changes in Krebs von den Lungen-6 levels in COVID-19 patients with different types in China. *Biosci Trends* 2020;14:290–6.
17 Bergantini L, Bargagli E, d'Alessandro M, *et al*. Prognostic Bioindicators in severe COVID-19 patients. *Cytokine* 2021;141:155455.
18 Arnold DT, Donald C, Lyon M, *et al*. Krebs von DEN Lungen 6 (KL-6) as a marker for disease severity and persistent radiological abnormalities following COVID-19 infection at 12 weeks. *PLOS ONE* 2021;16:e0249607.
19 Frix AN, Schoneveld L, Ladang A, *et al*. Could KL-6 levels in COVID-19 help to predict lung disease *Respir Res* 2020;21:309.
20 Castellví I, Castillo D, Corominas H, *et al*. Krebs von DEN Lungen-6 glycoprotein circulating levels are not useful as Prognostic marker in COVID-19 pneumonia: A large prospective cohort study. *Front Med (Lausanne)* 2022;9:973918.
21 Horimasu Y, Hattori N, Ishikawa N, *et al*. Different MUC1 gene polymorphisms in German and Japanese ethnicities affect serum KL-6 levels. *Respir Med* 2012;106:1756–64.
22 Janssen R, Kruit A, Grutters JC, *et al*. The Mucin-1 568 adenosine to guanine polymorphism influences serum Krebs von den Lungen-6 levels. *Am J Respir Cell Mol Biol* 2006;34:496–9.
23 Stock CJW, Hoyles RK, Daccord C, *et al*. Serum markers of pulmonary epithelial damage in systemic sclerosis-associated interstitial lung disease and disease progression. *Respirology* 2021;26:461–8.
24 Bonella F, Long X, Ohshimo S, *et al*. MUC1 gene polymorphisms are associated with serum KL-6 levels and pulmonary dysfunction in pulmonary alveolar Proteinosis. *Orphanet J Rare Dis* 2016;11:48.
25 Remuzgo-Martínez S, Atienza-Mateo B, Ocejo-Vinyals JG, *et al*. Role of MUC1 Rs4072037 polymorphism and serum KL-6 levels in patients with Antisynthetase syndrome. *Sci Rep* 2021;11:22574.
26 Namkoong H, Edahiro R, Takano T, *et al*. DOCK2 is involved in the host Genetics and biology of severe COVID-19. *Nature* 2022;609:754–60.
27 Tanaka H, Lee H, Morita A, *et al*. Clinical characteristics of patients with Coronavirus disease (COVID-19): preliminary baseline report of Japan COVID-19 task force, a nationwide consortium to investigate host genetics of COVID-19. *Int J Infect Dis* 2021;113:74–81.
28 Fukushima T, Chubachi S, Namkoong H, *et al*. Clinical significance of prediabetes, undiagnosed diabetes and diagnosed diabetes on critical outcomes in COVID-19: integrative analysis from the Japan COVID-19 task force. *Diabetes Obes Metab* 2023;25:144–55.
29 Otake S, Chubachi S, Namkoong H, *et al*. Clinical clustering with prognostic implications in Japanese COVID-19 patients: report from Japan COVID-19 task force, a nation-wide consortium to investigate COVID-19 host genetics. *BMC Infect Dis* 2022;22:735.
30 Fukushima T, Chubachi S, Namkoong H, *et al*. U-shaped association between abnormal serum uric acid levels and COVID-19 severity: reports from the Japan COVID-19 task force. *Int J Infect Dis* 2022;122:747–54.
31 Lee H, Chubachi S, Namkoong H, *et al*. Effects of mild obesity on outcomes in Japanese patients with COVID-19: a nationwide consortium to investigate COVID-19 host genetics. *Nutr Diabetes* 2022;12:38.
32 Ohtsuki Y, Kuroda N, Umeoka T, *et al*. KL-6 is another useful marker in assessing a micropapillary pattern in carcinomas of the breast and urinary bladder, but not the colon. *Med Mol Morphol* 2009;42:123–7.
33 Wu Z, McGoogan JM. Characteristics of and important lessons from the Coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese center for disease control and prevention. *JAMA* 2020;323:1239–42.
34 Docherty AB, Harrison EM, Green CA, *et al*. Features of 20 133 UK patients in hospital with COVID-19 using the ISARIC WHO clinical

characterisation protocol: prospective observational cohort study. *BMJ* 2020:m1985.
35 Deng G, Yin M, Chen X, *et al*. Clinical determinants for fatality of 44,672 patients with COVID-19. *Crit Care* 2020;24:179.
36 Harrison SL, Buckley BJR, Rivera-Caravaca JM, *et al*. Cardiovascular risk factors, cardiovascular disease, and COVID-19: an umbrella review of systematic reviews. *Eur Heart J Qual Care Clin Outcomes* 2021;7:330–9.
37 Shakaroun DA, Lazar MH, Horowitz JC, *et al*. Serum Ferritin as a Predictor of outcomes in hospitalized patients with COVID-19 pneumonia. *J Intensive Care Med* 2023;38:21–6.
38 Huang I, Pranata R, Lim MA, *et al*. C-reactive protein, procalcitonin, D-dimer, and ferritin in severe Coronavirus disease-2019: a meta-analysis. *Ther Adv Respir Dis* 2020;14:1753466620937175.
39 Wang G, Wu C, Zhang Q, *et al*. C-reactive protein level may predict the risk of COVID-19 aggravation. *Open Forum Infect Dis* 2020;7:ofaa153.
40 Xue M, Zhang T, Chen H, *et al*. Krebs von den Lungen-6 as a predictive indicator for the risk of secondary pulmonary fibrosis and its reversibility in COVID-19 patients. *Int J Biol Sci* 2021;17:1565–73.
41 Maruyama S, Nakamori Y, Nakano H, *et al*. Peak value of serum KL-6 may be useful for predicting poor prognosis of severe COVID-19 patients. *Eur J Med Res* 2022;27:69.
42 Katsura H, Sontake V, Tata A, *et al*. Human lung stem cell-based alveolospheres provide insights into SARS-CoV-2-mediated interferon responses and pneumocyte dysfunction. *Cell Stem Cell* 2020;27:890–904.
43 Lee JS, Lee EY, Ha Y-J, *et al*. Serum KL-6 levels reflect the severity of interstitial lung disease associated with connective tissue disease. *Arthritis Res Ther* 2019;21:58.
44 Yanaba K, Hasegawa M, Hamaguchi Y, *et al*. Longitudinal analysis of serum KL-6 levels in patients with systemic sclerosis: association with the activity of pulmonary fibrosis. *Clin Exp Rheumatol* 2003;21:429–36.
45 Sato H, Callister MEJ, Mumby S, *et al*. KL-6 levels are elevated in plasma from patients with acute respiratory distress syndrome. *Eur Respir J* 2004;23:142–5.
46 Fox SE, Akmatbekov A, Harbert JL, *et al*. Pulmonary and cardiac pathology in African American patients with COVID-19: an autopsy series from New Orleans. *Lancet Respir Med* 2020;8:681–6.
47 Adachi T, Chong J-M, Nakajima N, *et al*. Clinicopathologic and immunohistochemical findings from autopsy of patient with COVID-19, Japan. *Emerg Infect Dis* 2020;26:2157–61.
48 Kondo T, Hattori N, Ishikawa N, *et al*. KL-6 concentration in pulmonary epithelial lining fluid is a useful prognostic indicator in patients with acute respiratory distress syndrome. *Respir Res* 2011;12:32.
49 Zeng H-L, Chen D, Yan J, *et al*. Proteomic characteristics of bronchoalveolar lavage fluid in critical COVID-19 patients. *FEBS J* 2021;288:5190–200.
50 Varble N, Blain M, Kassin M, *et al*. Correction to: CT and clinical assessment in asymptomatic and pre-symptomatic patients with early SARS-CoV-2 in outbreak settings. *Eur Radiol* 2021;31:4406.
51 Yokoyama A, Kondo K, Nakajima M, *et al*. Prognostic value of circulating KL-6 in idiopathic pulmonary fibrosis. *Respirology* 2006;11:164–8.
52 Imbert Y, Foulks GN, Brennan MD, *et al*. Muc1 and estrogen receptor alpha gene polymorphisms in dry eye patients. *Exp Eye Res* 2009;88:334–8.
53 Ng W, Loh AXW, Teixeira AS, *et al*. Genetic regulation of MUC1 alternative splicing in human tissues. *Br J Cancer* 2008;99:978–85.
54 Kyo Y, Kato K, Park YS, *et al*. Antiinflammatory role of MUC1 mucin during infection with nontypeable haemophilus influenzae. *Am J Respir Cell Mol Biol* 2012;46:149–56.
55 Li Y, Dinwiddie DL, Harrod KS, *et al*. Anti-inflammatory effect of MUC1 during respiratory syncytial virus infection of lung epithelial cells in vitro. *Am J Physiol Lung Cell Mol Physiol* 2010;298:L558–63.
56 Michelen M, Manoharan L, Elkheir N, *et al*. Characterising long COVID: a living systematic review. *BMJ Glob Health* 2021;6:e005427.
57 Nalbandian A, Sehgal K, Gupta A, *et al*. Post-acute COVID-19 syndrome. *Nat Med* 2021;27:601–15.
58 Chopra V, Flanders SA, O'Malley M, *et al*. Sixty-day outcomes among patients hospitalized with COVID-19. *Ann Intern Med* 2021;174:576–8.
59 So M, Kabata H, Fukunaga K, *et al*. Radiological and functional lung sequelae of COVID-19: a systematic review and meta-analysis. *BMC Pulm Med* 2021;21:97.
60 Huang C, Huang L, Wang Y, *et al*. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* 2021;397:220–32.
61 Anastasi E, Manganaro L, Guiducci E, *et al*. Association of serum Krebs von den Lungen-6 and chest CT as potential prognostic factors in severe acute respiratory syndrome SARS-Cov-2: a preliminary experience. *Radiol Med* 2022;127:725–32.
62 Deng K, Fan Q, Yang Y, *et al*. Prognostic roles of KL-6 in disease severity and lung injury in COVID-19 patients: a longitudinal retrospective analysis. *J Med Virol* 2021;93:2505–12.
63 Brasen CL, Christensen H, Olsen DA, *et al*. Daily monitoring of viral load measured as SARS-Cov-2 antigen and RNA in blood, IL-6, CRP and complement C3D predicts outcome in patients hospitalized with COVID-19. *Clin Chem Lab Med* 2021;59:1988–97.
64 Yamaya T, Hagiwara E, Baba T, *et al*. Serum Krebs von den Lungen-6 levels are associated with mortality and severity in patients with Coronavirus disease 2019. *Respir Investig* 2021;59:596–601.