scientific reports

Check for updates

OPEN

Longitudinal changes in KL6 levels predict acute exacerbation in fibrotic interstitial lung disease

Akihiro Shiroshita[1,2,3]✉, Keisuke Anan[2,4], Chigusa Shirakawa[5], Hiroshi Shiba[6] & Nobuyuki Yajima[3]

Krebs von den Lungen-6 (KL-6) levels are potentially indicative markers of prospective acute exacerbation (AE) in interstitial lung disease (ILD); however, their longitudinal trends have not been sufficiently investigated. We investigated the predictive ability of patient-specific changes in serum KL-6 levels for predicting AE in patients with fibrotic ILD. We included patients with fibrotic ILDs from the RWD database in Japan who received antifibrotic therapy and had at least two serum KL-6 values during the follow-up. The outcome was AE defined based on primary diagnoses, emergency admission, and pulse/high-dose steroids on the day of or after admission. We used the joint regression model integrating longitudinal and survival analyses to assess the predictive ability of each patient's serial serum KL-6 measurements for AE. Among 939 patients with fibrotic ILDs, 194 (21%) patients experienced AE during follow-up (event rate, 0.13/person-year; 1-year incidence, 35%). The AE hazard ratio comparing patients with differing cumulative serum KL-6 levels was 1.54 (95% confidence interval: 1.20–1.98, $p < 0.001$). In conclusion, AE in fibrotic ILD patients on antifibrotic therapy may be predicted by high baseline KL-6 levels and their increasing trend. Serial KL-6 monitoring can serve as a valuable tool in a multifaceted approach.

Keywords Acute exacerbation, Fibrotic interstitial lung disease, Interstitial lung disease exacerbation, Interstitial pneumonia, KL-6, Serum Krebs von Den Lungen-6

Fibrotic interstitial lung disease (ILD) is a heterogeneous group of ILDs characterised by inflammation and fibrosis of the lung parenchyma[1]. It includes idiopathic pulmonary fibrosis (IPF) and other categories of idiopathic interstitial pneumonia, unclassifiable ILD, connective tissue disease-associated ILD (CTD-ILD), fibrotic hypersensitivity pneumonitis, sarcoidosis, organising pneumonia, drug-induced interstitial lung disorders, and ILD associated with occupational exposures. Approximately 20–30% of patients with ILD have fibrotic ILD[2–4]. These patients usually experience progressive fibrosis accompanied by worsening lung function, dyspnoea, poor physical performance, and poor quality of life[5].

Acute exacerbation (AE) of ILD is a life-threatening risk factor for subsequent disease progression[6,7]. Despite its clinical importance, no reliable markers have been established to predict future AEs. Previous studies have primarily evaluated baseline patient characteristics, such as dyspnoea, quality of life, body mass index, oxygenation, serum biomarkers, and pulmonary function tests[8–12]. However, baseline characteristics are only one aspect of the disease; as time elapses, additional information regarding disease progression becomes accessible. Recent declines in forced vital capacity and diffusing capacity of the lung for carbon monoxide may predict future AE and disease progression in IPF[8,9,11]. Nonetheless, some patients, such as older patients with cognitive impairment, apraxia, or hypoxia may have difficulty undergoing pulmonary function tests[13,14].

Krebs von den Lungen-6 (KL-6) is a high-molecular-weight glycoprotein, expressed by type II and bronchial epithelial cells[15]. Serum KL-6 levels can be assessed through a simple blood test, with results typically available within an hour. KL-6 was originally identified in Japan as a cancer biomarker and has subsequently been used for ILD diagnosis[15,16]. Recent studies suggested that both baseline levels and longitudinal trajectories in serum KL-6 might serve as useful predictors of future AE and disease progression in ILDs[17,18]. In particular, KL-6 may be a useful marker in fibrotic ILDs, as it plays a key role in the intra-alveolar fibrotic process and pulmonary

[1]Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, 2525 West End Avenue, Nashville, TN 37203, USA. [2]Scientific Research Works Peer Support Group (SRWS-PSG), Osaka, Japan. [3]Division of Rheumatology, Department of Medicine, Showa University School of Medicine, Tokyo, Japan. [4]Division of Respiratory Medicine, Saiseikai Kumamoto Hospital, Kumamoto, Japan. [5]Department of Respiratory Medicine, Kobe City Medical Center General Hospital, Hyogo, Japan. [6]Department of Internal Medicine, Suwa Central Hospital, Nagano, Japan. ✉email: akihiro.shiroshita@vanderbilt.edu

fibrosis[19]. However, the real-world evidence regarding the longitudinal role of serum KL-6 levels in predicting AE remains limited, mainly because the timing of KL-6 measurement was neither fixed nor random, and the presence and absence of KL-6 measurements may themselves reflect poor outcomes. This leads to a fundamental violation of the missing data assumption in many statistical models[20]. To address this limitation, our study aimed to leverage large-scale real-world data in Japan and evaluate the association between estimated patient-specific changes in serum KL-6 levels and the risk of AE in patients with fibrotic ILD. As many Japanese clinicians routinely measure serum KL-6 in clinical practice, both at baseline and during follow-up, since its insurance coverage in 1999, this dataset provides a unique opportunity to examine longitudinal patterns in a real-world setting.

Methods

Study design

This database research utilised the RWD database), anonymised commercial data maintained by the Health, Clinic, and Education Information Evaluation Institute (Kyoto, Japan) with support from JMDC Inc. (Tokyo, Japan). The data include electronic medical records (e.g., demographic, pharmacy, and laboratory data) and administrative claims information (e.g., disease name and procedure data) of approximately 20 million patients from over 200 medical institutions in Japan. This study adhered to the Declaration of Helsinki[21] and the REporting of studies Conducted using Observational Routinely-collected Data statements (Supplementary Table S1)[22]. The institutional review board of Showa University approved this study (approval number: 2023-119-B). The requirement for written informed consent was waived by the institutional review board of Showa University because the dataset had been de-identified.

Patient selection

The target population comprised patients with fibrotic ILD, aligning with the pathobiological context in which KL-6 is most mechanistically and clinically relevant as a potential biomarker. In addition, to enhance the specificity of identifying patients with fibrotic ILD, we restricted the cohort to those who were prescribed antifibrotic therapy. First, we selected patients with acute or chronic ILDs based on the 10th Revision of the International Statistical Classification of Diseases and Related Health Problems, including IPF, other idiopathic interstitial pneumonia, CTD-ILD, hypersensitivity pneumonitis, occupational lung diseases, drug-induced lung injury, radiation pneumonitis, Langerhans cell histiocytosis, pulmonary alveolar proteinosis, and sarcoidosis (Supplementary Table S2). Second, only patients who received antifibrotic therapy were included. Nintedanib and pirfenidone (Anatomical Therapeutic Chemical [ATC] Classification of L01EX09 and L04AX05) have shown therapeutic benefits in IPF and progressive fibrotic ILDs other than IPF (progressive fibrosing ILD)[6,23]. In Japan, nintedanib was approved for IPF treatment in 2015, ILD caused by systemic sclerosis in 2019, and progressive fibrosing ILDs in 2020. Pirfenidone was approved for treating IPF in 2008. In the current study, the first day of antifibrotic therapy was set as the index date. From the index date, patients at the same medical institute were longitudinally followed up until AE of fibrotic ILD or censoring (Fig. 1). Individuals with serum KL-6 levels measured once or less time from the index date to the end of the follow-up and those diagnosed

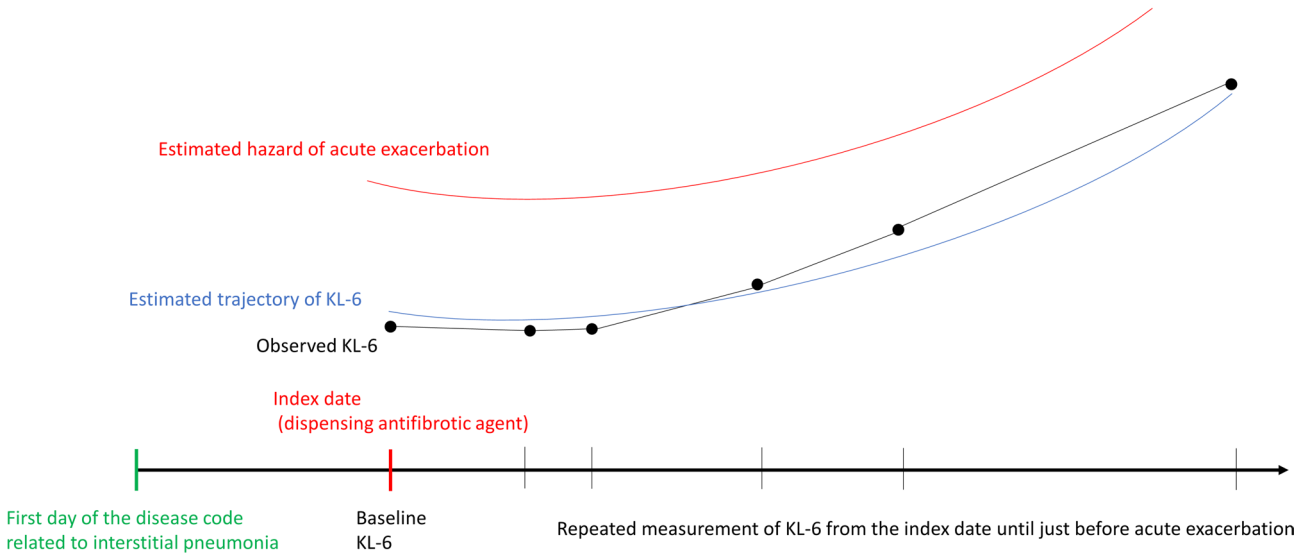

Fig. 1. Study overview. The index date referred to when a patient with a disease code related to interstitial pneumonia was prescribed anti-fibrotic therapy. The patient was tracked until acute exacerbation or censoring, and serial results of serum Krebs von den Lungen-6 levels until the end of follow-up were collected. The changes in serum KL-6 levels were modelled by longitudinal analysis, and the association with the hazard of acute exacerbation was modelled with survival analysis through a joint regression model.

with malignancy within 1 year before the index date were excluded. Patients with missing covariates were also excluded from the model.

Data extraction

The definitions of the variables are summarised in Supplementary Table S2. The following baseline patient characteristics were extracted from the electronic medical record data: age (continuous), sex (male vs. female), the disease name used for dispensing antifibrotic therapy, home oxygen therapy within 6 months before the index date, and the daily use of systemic steroids and immunosuppressive agents. The baseline body mass index, dyspnoea score (Hugh–Johns classification), activities of daily living (Barthel index), and comorbidity at the time of admission were extracted from the Yoshiki 1 file submitted to the government for reimbursement of in-hospital medical fees. Information on the Yoshiki 1 file was collected if patient admission occurred within 12 months prior to the index date. We calculated the Charlson Comorbidity Index based on the International Classification of Diseases 10th Revision Codes of Comorbidities (Supplementary Table S2).

Laboratory data were extracted from each hospital's database. The enzyme immunoassay, electrochemical luminescence immune assay, and latex agglutination tests have been approved for measuring serum KL-6 levels in Japan. All serum KL-6 measurements during the follow-up period were extracted, and the number of days from the index date to each measurement was calculated.

Outcome

The outcome of interest was the AE of fibrotic ILD. AE was defined based on combining the primary diagnoses related to AE of ILD (acute or chronic ILDs), emergency admissions, and the administration of pulse/high-dose steroids (≥ 50 mg/day of methylprednisolone equivalent) either on the admission day or the subsequent day (Supplementary Table S2)[24]. For the sensitivity analysis, we used other definitions of AE previously validated for AE of interstitial pneumonia and acute ILDs: one emphasised the specificity of AE, and the other emphasised the sensitivity by using only disease codes of AE without medication uses (Supplementary Table S2)[24]. These algorithms were validated by respiratory physicians using clinical information, including symptoms and high-resolution computed tomography findings as criteria for the reference standard.

Statistical analysis

Patient characteristics were presented as frequencies and proportions for categorical variables and as means with standard deviations or median and interquartile range, as appropriate, for continuous variables. The patient-specific trend of serum KL-6 was depicted in a spaghetti plot, accompanied by a smoothing line representing the average trend. A Kaplan–Meier plot was used to display survival probability over time. As an exploratory analysis, we illustrated spaghetti plots stratified by subgroups of background fibrotic ILDs, which were defined based on ICD-10 codes at the time of antifibrotic therapy initiation.

To evaluate the association between patient-specific changes in serum KL-6 levels and AE risk, we used the joint regression model involving the development of a linear mixed-effects model for serial serum KL-6 measurements in each patient and a proportional hazard regression model for the risk of AE[25]. The model could also accounted for measurement errors inherent in KL-6 assessments—the linear mixed-effects model aimed to estimate the underlying trajectory of KL-6 levels. The model included both time-independent covariates (age, sex, season of the index date, systemic steroid use, and immunosuppressive use at baseline) and a time-dependent covariate (days from the index date to each KL-6 measurement). Patient-specific random effects were incorporated into the intercept, as well as the number of days from the index date to each KL-6 measurement. The proportional hazards regression model was postulated using a Weibull baseline risk function. The same covariates as in the longitudinal model were incorporated, including the history of serum KL-6 defined within the linear mixed-effects model. The adaptive Gaussian–Hermite rule was applied for numerical integration within the maximum likelihood function. Given the right-skewed and non-negative distribution of serum KL-6, it was divided by 100 and log-transformed. The detailed information is described Supplementary Table S3.

We demonstrated the dynamic predictions of conditional survival using the joint regression model. Four extreme cases were simulated: a 65-year-old man not taking baseline steroids or immunosuppressants, initiating antifibrotic therapy in winter, with KL-6 measurements of (1) 500 U/mL at day 0, 1000 U/mL at day 300, and 2000 U/mL at day 600 (increasing trend); (2) 2000 U/mL at day 0, 1000 U/mL at day 300, and 500 U/mL at day 600 (decreasing trend); (3) 3000 U/mL at day 0, 3000 U/mL at day 300, and 3000 U/mL at day 600 (constantly high); and (4) 500 U/mL at day 0, 500 U/mL at day 300, and 500 U/mL at day 600 (constantly low). Statistical significance was set at a two-sided p-value of < 0.05. We used R software version 4.4.2 (R Foundation for Statistical Computing, Vienna, Austria) for statistical analyses.

Results

Descriptive analysis

Figure 2 summarises the patient selection process. We identified 41,658 patients with disease codes related to fibrotic ILD; among them, 1,112 patients were prescribed antifibrotic therapy. After excluding 173 patients with ≤ 1 measurement of KL-6, 939 patients with fibrotic ILDs were included in this study. Tables 1 and 2 summarise the baseline patient characteristics. A total of 312/939 (33%) patients had IPF, and 113/939 (12%) patients had CTD-ILD (rheumatoid arthritis, 4%; systemic sclerosis, 4%). Among the included patients, 194 (21%) experienced AE during the follow-up period (event rate: 0.13/person-year; 1-year incidence: 35%). Among patients without AE, 86 (12%) patients died from other causes within the first year of follow-up. No patients developed malignancy before the index date and during the follow-up period. After the index date, 96% of patients continued antifibrotic therapy without a gap of one month or more until the time of their first acute exacerbation (mean treatment duration: 1.4 years).

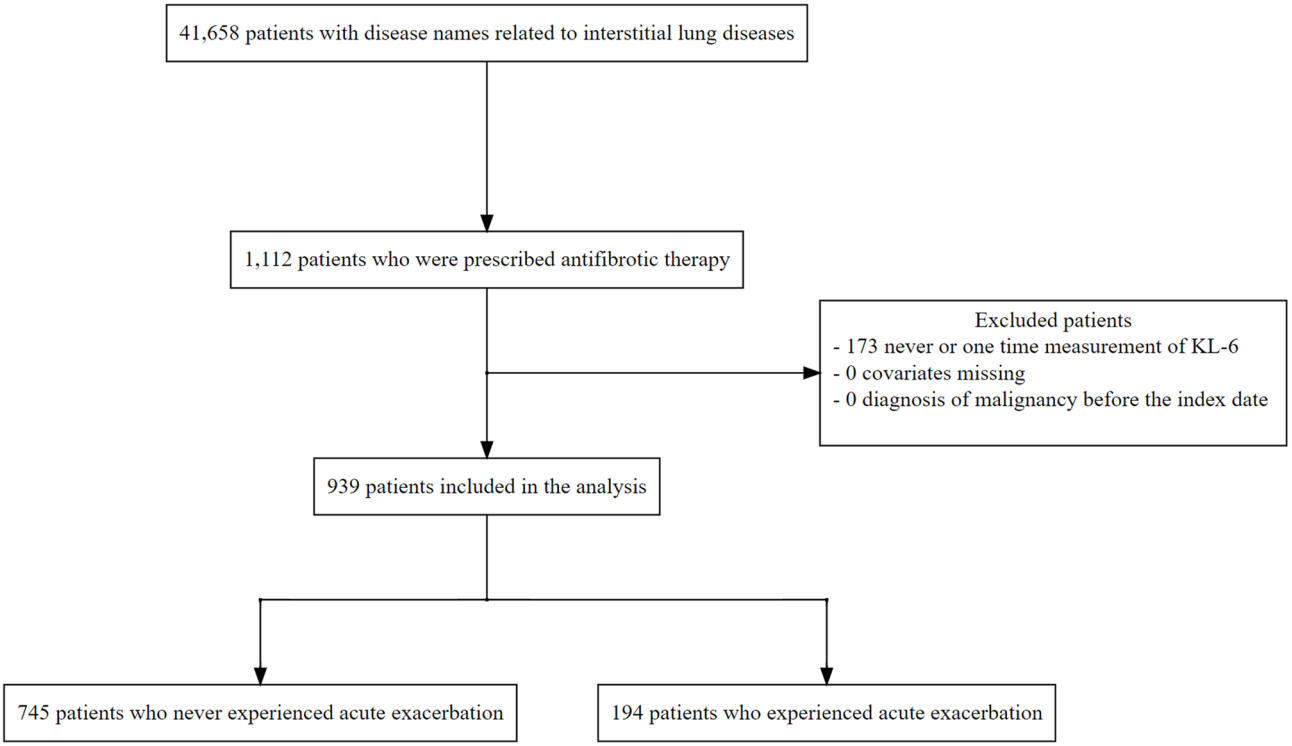

Fig. 2. Flow diagram of patient selection. Overall, 41,658 patients with a disease code of interstitial lung disease (ILD) were identified; among them, 1,112 patients received antifibrotic therapy. After excluding 173 patients with ≤ 1 measurement of Krebs von den Lungen-6 levels, 939 patients with fibrotic ILDs were included in the final analysis.

The baseline serum KL-6 levels were as follows: minimum, 265 U/mL; first quartile, 735 U/mL; median, 1,134 U/mL; mean, 1,406 U/mL; third quartile, 1,680 U/mL; and maximum, 12,408 U/mL. KL-6 measurements were performed at 3–170 visits (median, 11; interquartile range, 6–20) from the index date until the end of the follow-up. The difference between two consecutive measurements varied widely (minimum: −4750 U/mL; first quartile, −94 U/mL; median: −11 U/mL; mean: −16 U/mL; third quartile, 58 U/mL; maximum: 6,730 U/mL). Figure 3a shows the individual-level trajectories of KL-6 from the index date, accompanied by a smoothing line representing the average trend. Patients who experienced AEs had higher serum KL-6 levels over time than those without AE. This overall pattern was consistent across subgroup-specific spaghetti plots, although the trend in KL-6 levels among patients with IPF showed a steeper rise compared to that observed in patients with connective tissue disease–associated ILD (Supplementary Figure S1). Figure 3b shows the individual-level trajectories of serum KL-6 at the end of the follow-up. A slight increasing trend in serum KL-6 levels was observed before AE, whereas the levels before censoring remained relatively stable. The Kaplan–Meier curve showed a constant or slightly monotonic increase in AE risk over time (Supplementary Figure S2).

Main analysis

We observed a strong association between the log of serum KL-6 levels and AE occurrence over time (Table 3). The hazard ratio for AE comparing patients with a cumulative serum KL-6 difference of 100*exp(1) = 272 U/mL was 1.54 (95% confidence interval: 1.20–1.98, $p < 0.001$). As shown in Supplementary Figure S3, patients with an increasing trend of serum KL-6 levels and constantly high serum KL-6 levels had a more unfavourable predicted survival compared with those with a decreasing trend or consistently low serum KL-6 values. Sensitivity analysis using alternate algorithms for detecting AE yielded similar results (sensitivity analysis 1: hazard ratio = 1.26, 95% confidence interval = 1.01–1.56, $p = 0.04$, sensitivity analysis 2, hazard ratio = 1.30, 95% confidence interval = 1.16–1.45, $p < 0.001$).

Discussion

Only limited evidence exists regarding the dynamic prediction of AE. To the best of our knowledge, this study is the first to demonstrate a robust capability of patient-specific changes in serum KL-6 levels for predicting AE in patients with fibrotic ILD. The trajectories of serum KL-6 exhibited high heterogeneity and fluctuation; however, physicians should be vigilant when observing an increasing trend in KL-6 or persistently elevated serum KL-6 values, as they may indicate prospective AE. Serum KL-6 monitoring can be used as a component of a multifaceted approach to risk assessment.

Previous studies primarily evaluated baseline patient characteristics, identifying sex, age, blood biomarkers, pulmonary function tests, and radiological and histological patterns as potential predictors of poor prognosis,

	Without acute exacerbation (*N* = 745)	With acute exacerbation (*N* = 194)	Total (*N* = 939)
Age (year), mean (SD)	72 (8)	71 (9)	72 (8)
Male	517 (69)	143 (74)	660 (70)
Nintedanib	473 (64)	104 (54)	577 (61)
Baseline KL-6 level, median (IQR)	1076 (686–1619)	1172 (840–1745)	1100 (710–1623)
Season of the index date			
Winter	200 (27)	51 (26)	251 (27)
Spring	202 (27)	40 (21)	242 (26)
Summer	161 (22)	53 (27)	214 (23)
Autumn	182 (24)	50 (26)	232 (25)
BMI			
<18.5	34 (5)	5 (3)	39 (4)
18.5 to < 25	179 (24)	50 (26)	229 (24)
≥25	102 (14)	44 (23)	146 (16)
Missing	430 (58)	95 (49)	525 (56)
Activity of daily living (Barthel index)			
<20	8 (1)	1 (1)	9 (1)
20 to < 50	6 (1)	5 (3)	11 (1.2)
>50	310 (42)	95 (49)	405 (43)
Missing	421 (57)	93 (48)	514 (55)
Charlson Comorbidity Index (median, IQR)	3 (2–3)	3 (2–3)	3 (2–3)
Home oxygen therapy	118 (16)	38 (20)	156 (17)
Dyspnoea score (Hugh–Johns classification)			
<3	206 (28)	58 (30)	264 (28)
≥3	72 (10)	34 (18)	106 (11)
Missing	467 (63)	102 (53)	569 (61)
Daily systemic steroids	135 (18)	53 (27)	188 (20)
Daily immunosuppressive agents	29 (4)	16 (8)	45 (5)

Table 1. Patient characteristics at baseline. Mechanical ventilation included invasive/non-invasive ventilation, high-flow nasal cannula, and continuous positive airway pressure. N, number; SD, standard deviation; KL-6, Krebs von den Lungen-6; IQR, interquartile range; BMI, body mass index.

including AE[2,26–29]. To date, no single established marker has been identified, as disease processes are driven by complex, integrated mechanisms. Our study focused on serum KL-6 levels, a simple and rapid blood test with biological plausibility in fibrotic ILDs. Unlike many prior studies that emphasized baseline values, we demonstrated that serum KL-6 levels exhibit considerable fluctuation over time and that the prognostic value of baseline KL-6 is influenced by the timing of fibrotic ILD diagnosis. These findings highlight the importance of evaluating longitudinal trends rather than relying solely on baseline measurements.

Although some previous studies have explored serial KL-6 trajectories in relation to poor outcomes, they faced notable methodological challenges[18,30–32]. For example, they evaluated the correlation between serial KL-6 levels and patient outcomes but excluded patients when either KL-6 or outcomes were missing. This exclusion strategy assumed missing-completely-at-random; the probability of an outcome being missing was independent of both observed and unobserved factors[20]. However, in real-world clinical settings without standardised measurement protocols, missingness is often informative; for example, clinical deterioration may lead to missed outcomes. Furthermore, most of these prior studies had limited sample sizes and were unable to evaluate clinically meaningful endpoints such as AE, relying instead on surrogate outcomes like pulmonary function test results. This study directly addressed AE, and employed the joint regression model, which combines longitudinal data analysis for exposure with survival analysis for the outcome. This meant that we estimated the latent trend in true values of KL-6 over the follow-up period for each patient, allowing us to account for the informative timing of KL-6 measurements.

Our study offers several research implications. First, future studies with larger sample sizes could explore the serial trajectories of multiple biomarkers over time such as pulmonary function test results, dyspnoea scores, physical function, and image findings. Using the joint regression model, researchers can incorporate multiple time-dependent exposures and assess their predictive value in more realistic, real-world settings. Second, rather than treating fibrotic ILD as a single, homogeneous group, subgroup-specific models should be considered. As shown in Supplementary Figure S1, the trend in KL-6 levels among patients with IPF may differ from those with connective tissue disease–associated ILD, suggesting that distinct predictive patterns may exist across subtypes. Finally, expanding the disease scope to include non-fibrotic ILDs could further broaden the clinical applicability of KL-6 as a biomarker in the management of ILD.

	Number of patients (%)
Idiopathic interstitial pneumonia	
Idiopathic pulmonary fibrosis	312 (32.2)
Usual interstitial pneumonia	5 (0.5)
Nonspecific interstitial pneumonia	15 (1.6)
Combined pulmonary fibrosis and Emphysema	10 (1.1)
Cryptogenic organising pneumonia	1 (0.1)
Idiopathic interstitial pneumonia (unspecified)	156 (16.6)
Connective tissue disease-associated ILD	
Rheumatoid arthritis	39 (4.2)
Systemic sclerosis	37 (3.9)
Sjogren syndrome	16 (1.7)
Microscopic polyangiitis	7 (0.7)
Connective tissue disease-associated ILD (unspecified)	9 (1.0)
Systemic lupus erythematosus	3 (0.3)
IgG4-related disease	2 (0.2)
Hypersensitivity pneumonitis	15 (1.6)
Eosinophilic pneumonia	1 (0.1)
Others	
Interstitial pneumonia (unspecified)	261 (27.8)
Lung fibrosis (unspecified)	25 (2.7)
Diffuse interstitial pneumonia (unspecified)	17 (1.8)
Acute interstitial pneumonia (unspecified)	6 (0.6)
Pulmonary haemorrhage	2 (0.2)

Table 2. Background interstitial lung disease in patients. ILD, interstitial lung disease; IgG, immunoglobulin.

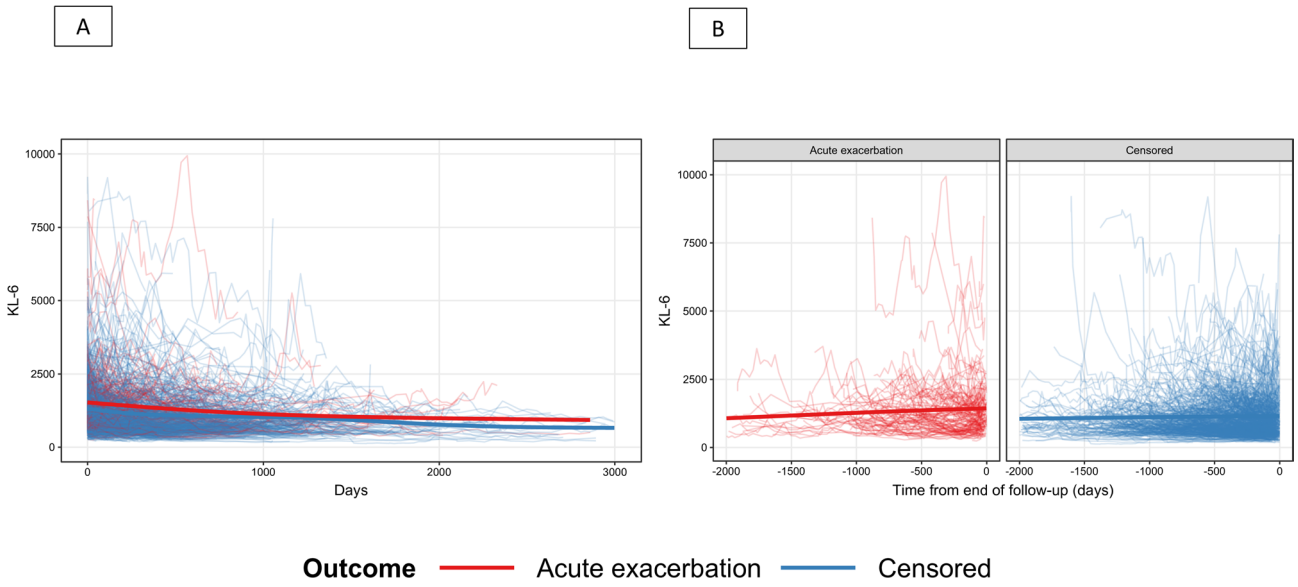

Fig. 3. Spaghetti plots of Krebs von den Lungen-6 (KL-6) levels. (**a**) Individual-level trajectories of KL-6 levels from the index date, along with a smoothing line for the average trend. On average, patients with acute exacerbation (AE) have higher serum KL-6 values over time than those without. (**b**) Individual-level trajectories of serum KL-6 from the end of the follow-up. A slight increasing trend in the serum KL-6 is observed before AE, and the trend before censoring is flat.

A caveat in interpreting the study results was that we primarily aimed to evaluate the predictive ability of KL-6 rather than identifying the independent association between KL-6 and AE[33]. In particular, we aimed to provide information patients for whom closer monitoring may be warranted based solely on KL-6 levels, irrespective of other biomarkers. In this context, potential confounding factors—such as pulmonary function, which may be associated with both KL-6 and AE—could have distorted the observed association. Consequently, our study

	Estimate	95% Confidence interval	*p*-value
Longitudinal process			
Days from the index date to KL-6 measurement	1.00	1.00–1.00	< 0.001
Baseline daily steroid use	1.12	1.01–1.23	0.031
Baseline daily immunosuppressant Use	0.98	0.81–1.18	0.824
Age	0.99	0.99–0.99	< 0.001
Sex	0.98	0.90–1.06	0.528
Season	0.99	0.96–1.03	0.642
Event process (hazard ratio)			
Baseline steroid	1.79	1.27–2.53	< 0.001
Baseline immunosuppressant	1.70	0.96–3.00	0.067
Age	1.02	1.00–1.04	0.062
Sex	1.49	1.07–2.09	0.018
Season	1.06	0.94–1.20	0.337
Association of log (KL-6/100)	1.54	1.19–1.98	< 0.001

Table 3. Summary of analysis results. KL-6, Krebs von den Lungen-6.

does not support a clinically meaningful KL-6 cut-off value, and the predictive performance we estimated is context-specific and requires external validation in other clinical settings.

This study has other limitations. First, outcome misclassification may have occurred, although we used modified versions of previously validated algorithms. Especially, given the complexity of distinguishing true AE from disease progression, further validation studies are warranted to more rigorously incorporate imaging findings as a criterion for AE identification. Second, in our database, longitudinal follow-ups were conducted only at the same hospital; this could have skewed associations toward null, potentially increasing the likelihood of missing AE more than overdiagnosing these events. Third, the small sample size limited our ability to develop a more intricate and flexible model. For instance, we could not evaluate the contribution of the rate of change in patient-specific KL-6 levels. Fourth, our study focused on the Japanese population. The genotype distribution may cause differences in KL-6 levels[34,35]. Further large-scale studies are needed to validate our findings and to develop a more accurate dynamic prediction model for AE. Lastly, we were unable to capture additional important treatments beyond what is available in administrative claims data, such as exposure removal in cases of hypersensitivity pneumonitis. This may have influenced the observed association between KL-6 levels and acute exacerbation and underscores the need for further validation studies in different settings.

In conclusion, high baseline KL-6 levels and an increasing trend can predict AE in fibrotic ILD. Physicians should consider longitudinal KL-6 measurements, as a component of a multifaceted approach to risk assessment. Our results should be validated in other settings.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request. The data are not publicly available because of the restrictions imposed by JMDC Inc.

Received: 18 October 2024; Accepted: 16 July 2025
Published online: 18 July 2025

References

1. Wong, A. W., Ryerson, C. J. & Guler, S. A. Progression of fibrosing interstitial lung disease. *Respir. Res.***21**32. (2020).
2. Zamora-Legoff, J. A., Krause, M. L., Crowson, C. S., Ryu, J. H. & Matteson, E. L. Progressive decline of lung function in rheumatoid arthritis-associated interstitial lung disease. *Arthritis Rheumatol.***69**, 542–549 (2017).
3. Reiseter, S. et al. Progression and mortality of interstitial lung disease in mixed connective tissue disease: A long-term observational nationwide cohort study. *Rheumatology***57**, 255–262 (2018).
4. Gagliardi, M. et al. Real-life prevalence of progressive fibrosing interstitial lung diseases. *Sci. Rep.***11**, 23988 (2021).
5. Cottin, V. et al. Fibrosing interstitial lung diseases: Knowns and unknowns. *Eur. Respir. Rev.***28**, 180100 (2019).
6. Raghu, G. et al. Idiopathic pulmonary fibrosis (an update) and progressive pulmonary fibrosis in adults: An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am. J. Respir. Crit. Care Med.***205**, e18–e47 (2022).
7. Collard, H. R. et al. Acute exacerbation of idiopathic pulmonary fibrosis. An international working group report. *Am. J. Respir. Crit. Care Med.***194**, 265–275 (2016).
8. Kondoh, Y. et al. Risk factors of acute exacerbation of idiopathic pulmonary fibrosis. *Sarcoidosis Vasc Diffuse Lung Dis.* **27**, 103–110 (2010).
9. Song, J. W., Hong, S. B., Lim, C. M., Koh, Y. & Kim, D. S. Acute exacerbation of idiopathic pulmonary fibrosis: Incidence, risk factors and outcome. *Eur. Respir. J.***37**, 356–363 (2011).
10. Alqalyoobi, S. et al. Circulating plasma biomarkers of progressive interstitial lung disease. *Am. J. Respir. Crit. Care Med.***201**, 250–253 (2020).
11. Collard, H. R. et al. Suspected acute exacerbation of idiopathic pulmonary fibrosis as an outcome measure in clinical trials. *Respir. Res.***14**, 73 (2013).
12. Ohshimo, S. et al. Baseline KL-6 predicts increased risk for acute exacerbation of idiopathic pulmonary fibrosis. *Respir. Med.***108**, 1031–1039 (2014).

13. Allen, S. C. & Baxter, M. A comparison of four tests of cognition as predictors of inability to perform spirometry in old age. *Age Ageing***38**, 537–541 (2009).
14. Carvalhaes-Neto, N. et al. Cognitive function and assessment of lung function in the elderly. *Am. J. Respir. Crit. Care Med.***152**, 1611–1615 (1995).
15. Kohno, N. et al. Detection of soluble tumor-associated antigens in Sera and effusions using novel monoclonal antibodies, KL-3 and KL-6, against lung adenocarcinoma. *Jpn J. Clin. Oncol.* **18**, 203–216 (1988).
16. Wang, C., Wang, Q., Liu, T., Zhu, J. & Zhang, B. Krebs von Den Lungen-6 (KL-6) as a diagnostic marker for pulmonary fibrosis: a systematic review and meta-analysis. *Clin. Biochem.* **114**, 30–38 (2023).
17. D'Alessandro, M. et al. Krebs von Den Lungen-6 as a biomarker for disease severity assessment in interstitial lung disease: a comprehensive review. *Biomark. Med.***14**, 665–674 (2020).
18. Watanabe, S. et al. Kinetic changes in serum KL-6 levels predict disease progression in patients with systemic sclerosis-associated interstitial lung disease. *Respir. Med.***191**, 106689 (2022).
19. Hirasawa, Y. et al. KL-6, a human MUC1 mucin, is chemotactic for human fibroblasts. *Am. J. Respir. Cell Mol. Biol.***17**, 501–507 (1997).
20. Rubin, D. B. Inference and missing data. *Biometrika* **63**, 581–592 (1976).
21. World Medical Association. World medical association declaration of helsinki. Ethical principles for medical research involving human subjects. *Bull. World Health Organ.***79**, 373–374 (2001).
22. Benchimol, E. I. et al. The reporting of studies conducted using observational routinely-collected health data (RECORD) statement. *PLoS Med.***12**, e1001885 (2015).
23. Behr, J. et al. Pirfenidone in patients with progressive fibrotic interstitial lung diseases other than idiopathic pulmonary fibrosis (RELIEF): a double-blind, randomised, placebo-controlled, phase 2b trial. *Lancet Respir Med.* **9**, 476–486 (2021).
24. Anan, K. et al. Algorithms identifying patients with acute exacerbation of interstitial pneumonia and acute interstitial lung diseases developed using Japanese administrative data. *Cureus* **16**, e53073 (2024).
25. Wulfsohn, M. S. & Tsiatis, A. A. A joint model for survival and longitudinal data measured with error. *Biometrics* **53**, 330–339 (1997).
26. Gimenez, A. et al. Change in FVC and survival in chronic fibrotic hypersensitivity pneumonitis. *Thorax* **73**, 391–392 (2018).
27. Ryerson, C. J. et al. Predicting survival across chronic interstitial lung disease. *Chest***145**, 723–728 (2014).
28. Qiu, M., Chen, Y. & Ye, Q. Risk factors for acute exacerbation of idiopathic pulmonary fibrosis: A systematic review and meta-analysis. *Clin. Respir. J.***12**, 1084–1092 (2018).
29. Zhang, T., Shen, P., Duan, C. & Gao, L. KL-6 as an immunological biomarker predicts the severity, progression, acute exacerbation, and poor outcomes of interstitial lung disease: A systematic review and meta-analysis. *Front. Immunol.***12**, 745233 (2021).
30. Yanaba, K. et al. Longitudinal analysis of serum KL-6 levels in patients with systemic sclerosis: association with the activity of pulmonary fibrosis. *Clin. Exp. Rheumatol.* **21**, 429–436 (2003).
31. Bergantini, L. et al. Serial KL-6 analysis in patients with idiopathic pulmonary fibrosis treated with nintedanib. *Respir Investig.* **57**, 290–291 (2019).
32. Wakamatsu, K. et al. Prognostic value of serial serum KL-6 measurements in patients with idiopathic pulmonary fibrosis. *Respir Investig.* **55**, 16–23 (2017).
33. Sainani, K. L. Explanatory versus predictive modeling. *PM R***6**, 841–844 (2014).
34. Horimasu, Y. et al. Different *MUC1* gene polymorphisms in German and Japanese ethnicities affect serum KL-6 levels. *Respir. Med.***106**, 1756–1764 (2012).
35. Bonella, F. et al. *MUC1* gene polymorphisms are associated with serum KL-6 levels and pulmonary dysfunction in pulmonary alveolar proteinosis. *Orphanet J. Rare Dis.***11**, 48 (2016).

Acknowledgements

The authors thanks to Health, Clinic, and Education Information Evaluation Institute and JMDC Inc, for the database development for the study.

Author contributions

AS, KA, CS, HS, and NY contributed to the conception and design of the study. AS managed the database and performed the data analysis. All authors confirmed the validity of the data analysis and contributed to the data interpretation. AS and KA drafted the manuscript. All authors critically revised and approved the final version of the manuscript. AS, KA, CS, HS, and NY consent to be accountable for all aspects of this study.

Funding

This study was funded by the Japan Society for the Promotion of Science and Pfizer Health Research Foundation. The Japanese Society for the Promotion of Science provided English editing and publication fees (Grants-in-Aid for Scientific Research [Kakenhi]; grant number 23K09582). The Pfizer Health Research Foundation supported the cost of data acquisition. The funders played no role in the study design, execution, analyses, interpretation, or decision to submit the results.

Declarations

Conflict of interest

AS received funding from the Japan Society for the Promotion of Science and Pfizer Health Research Foundation. Other authors do not have a conflict of interest.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-12236-8.

Correspondence and requests for materials should be addressed to A.S.