Title

Quantitative interstitial lung disease scores in idiopathic inflammatory myopathies: longitudinal changes and clinical implications.

Journal

Rheumatology, 62(11)

Authors

Yeo, Jina
Yoon, Soon
Kim, Ju
et al.

Publication Date

2023-11-02

DOI

10.1093/rheumatology/kead122

Peer reviewed

Rheumatology, 2023, **62**, 3690–3699
https://doi.org/10.1093/rheumatology/kead122
Advance access publication 17 March 2023
Original Article

British Society for Rheumatology RHEUMATOLOGY OXFORD

Clinical science

Quantitative interstitial lung disease scores in idiopathic inflammatory myopathies: longitudinal changes and clinical implications

Jina Yeo [1,‡], Soon Ho Yoon[2,‡], Ju Yeon Kim[3], Jeong Seok Lee[4,5], Eun Young Lee [3,6], Jin Mo Goo[2], Lila Pourzand[7], Jonathan G. Goldin[7], Grace-Hyun J. Kim [7,§], You-Jung Ha [6,8,*,§]

[1]Division of Rheumatology, Department of Internal Medicine, Gachon University Gil Medical Center, Incheon, Republic of Korea
[2]Department of Radiology, Seoul National University Hospital, Seoul National University College of Medicine, Seoul, Republic of Korea
[3]Division of Rheumatology, Department of Internal Medicine, Seoul National University Hospital, Seoul, Republic of Korea
[4]Clinic Pappalardo Center, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea
[5]GENOME INSIGHT Inc, Daejeon, Republic of Korea
[6]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea
[7]Department of Radiological Sciences, David-Geffen School of Medicine, University of California, Los Angeles, CA, USA
[8]Division of Rheumatology, Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam, Gyeonggi-Do, Republic of Korea

*Correspondence to: You-Jung Ha, Division of Rheumatology, Department of Internal Medicine, Seoul National University Bundang Hospital, 82 Gumi-ro, 173beon-gil, Bundang-gu, Seongnam, Gyeonggi-do 13620, Republic of Korea. E-mail: hayouya@snubh.org
[‡]J.Y. and S.H.Y. contributed equally.
[§]G.-H.J.K. and Y.-J.H. contributed equally.

Abstract

Objectives: To investigate computer-aided quantitative scores from high-resolution CT (HRCT) images and determine their longitudinal changes and clinical significance in patients with idiopathic inflammatory myopathies (IIMs)-related interstitial lung disease (IIMs-ILD).

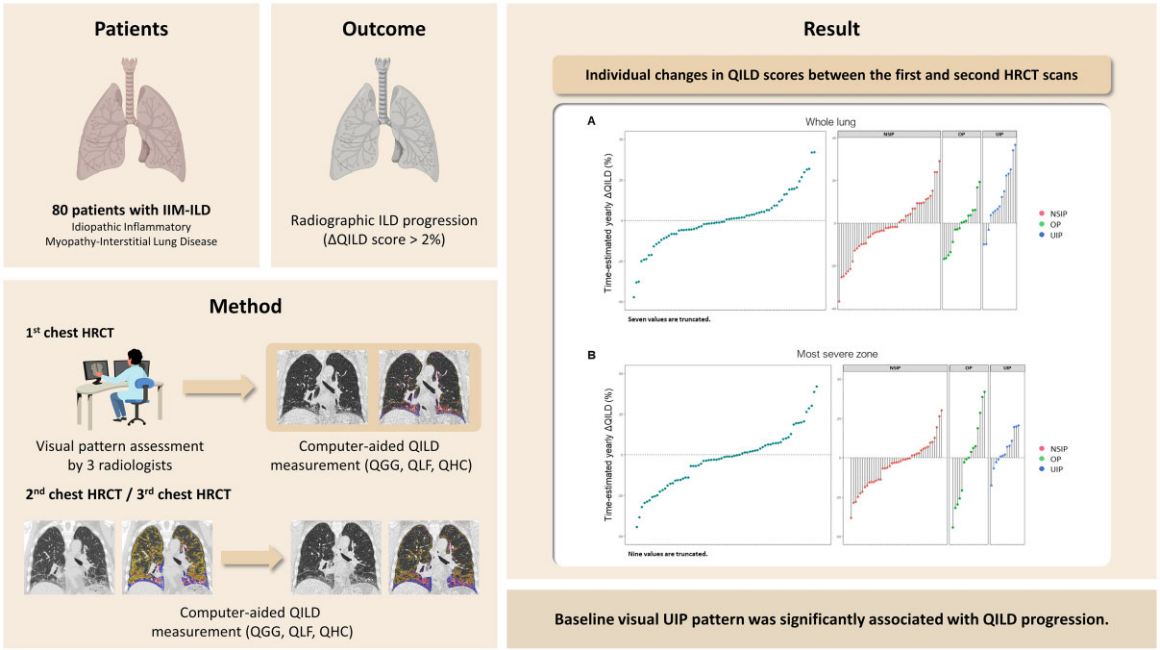
Methods: The clinical data and HRCT images of 80 patients with IIMs who underwent serial HRCT scans at least twice were retrospectively analysed. Quantitative ILD (QILD) scores (%) were calculated as the sum of the extent of lung fibrosis, ground-glass opacity, and honeycombing. The individual time-estimated ΔQILD between two consecutive scans was derived using a linear approximation of yearly changes.

Results: The baseline median QILD (interquartile range) scores in the whole lung were 28.1% (19.1–43.8). The QILD was significantly correlated with forced vital capacity ($r=-0.349$, $P=0.002$) and diffusing capacity for carbon monoxide ($r=-0.381$, $P=0.001$). For ΔQILD between the first two scans, according to the visual ILD subtype, QILD aggravation was more frequent in patients with usual interstitial pneumonia (UIP) than non-UIP (80.0% *vs* 44.4%, $P=0.013$). Multivariable logistic regression analyses identified UIP was significantly related to radiographic ILD progression (ΔQILD >2%, $P=0.015$). Patients with higher baseline QILD scores (>28.1%) had a higher risk of lung transplantation or death ($P=0.015$). In the analysis of three serial HRCT scans ($n=41$), dynamic ΔQILD with four distinct patterns (improving, worsening, convex and concave) was observed.

Conclusion: QILD changes in IIMs-ILD were dynamic, and baseline UIP patterns seemed to be related to a longitudinal progression in QILD. These may be potential imaging biomarkers for lung function, changes in ILD severity and prognosis in IIMs-ILD.

Received: 22 July 2022. **Accepted:** 3 March 2023

Graphical Abstract



Keywords: idiopathic inflammatory myopathy, interstitial lung disease, lung transplant-free survival, quantitative score

Rheumatology key messages

- Quantitative ILD (QILD) scores can be markers of lung function and longitudinal changes in IIMs-ILD.
- The baseline UIP pattern was related to radiographic ILD progression in IIMs-ILD.
- High baseline QILD scores may help differentiate a poor prognostic group of IIMs-ILD.

Introduction

Idiopathic inflammatory myopathies (IIMs) are a rare and heterogeneous group of autoimmune conditions characterized by chronic skeletal muscle inflammation [1]. IIMs include PM, DM, IBM, immune-mediated necrotizing myopathy, myositis of the anti-synthetase syndrome and myositis as part of the overlap syndrome. Extramuscular manifestations are frequently observed in IIMs, including in the skin, joints, heart, lung and gastrointestinal tract. IIMs-associated interstitial lung disease (IIMs-ILD) is a serious extramuscular manifestation that leads to a significant increase in morbidity and mortality [2, 3]. This may precede or follow the onset of muscle or skin involvement. The clinical course of IIMs-ILD is hypervariable, and some patients can experience rapidly progressive ILD, which is associated with a poor prognosis and a mortality rate of over 70% [4]. Treatment of IIMs-ILDs is based on high-dose glucocorticoids in combination with immunosuppressive drugs, including CYC, MMF, AZA, calcineurin inhibitors such as tacrolimus or ciclosporin, rituximab and IVIG. However, because of the absence of randomized clinical trials, its management is still empirical and remains challenging [5, 6].

High-resolution computed tomography (HRCT) is the best method for detecting and characterizing ILD as well as predicting outcomes [7]. In reports of other connective tissue disorder (CTD)-ILD, non-specific interstitial pneumonia (NSIP) is most commonly reported [6]. While usual interstitial pneumonia (UIP) in idiopathic interstitial pneumonia tends to respond poorly to glucocorticoids and immunosuppressants, UIP and NSIP have shown similar outcomes in CTD-ILD [8]. Moreover, HRCT parenchymal changes in IIMs-ILD are heterogeneous, and two different HRCT patterns may also coexist in the same patient. Several studies have suggested that ILD patterns based on HRCT can be related to disease course and treatment response, but the heterogeneity of IIMs-ILD makes an evaluation of the clinical course particularly difficult [9–12].

The quantitative ILD (QILD) score is a computer-based scoring system that assesses the total ILD burden by summation of specific features of HRCT parenchymal changes including lung fibrosis, ground-glass opacity and honeycombing, and it has been suggested as an objective tool to assess disease severity and prognosis in ILD [13–15]. For systemic sclerosis (SSc), the QILD score of HRCT has been comprehensively validated in assessing the severity of ILD [11]. Furthermore, the QILD score describes the quantitative and qualitative changes that are longitudinally related to immunosuppressive agents or clinical parameters used in SSc- or RA-related ILD [16, 17]. Some retrospective studies have suggested that possible prognostic information can be obtained

from baseline HRCT (e.g. radiologic visual pattern, specific zone involvement) in IIMs-ILD [18–21], but these results are not consensual, and no study has demonstrated QILD scores assessed on serial HRCT changes. In addition, the clinical impact of QILD scores and its longitudinal changes related to immunosuppressive agents or clinical parameters have not yet been evaluated in IIMs-ILD.

Thus, this study aimed to perform a comprehensive analysis of visual pattern assessment and QILD scores from serial HRCT images in patients with IIMs-ILD to investigate the longitudinal changes in QILD scores and to determine their possible relationships with clinical parameters, including prognosis.

Methods

Study population

We retrospectively evaluated patients with IIMs-ILD at two tertiary referral centres. We included patients with IIMs who underwent at least two serial thin-section HRCT scans and had matched pulmonary function parameter results between January 2006 and December 2017. Scans with an axial thickness of contiguous slices of >3 mm were excluded from the analysis. The Bohan and Peter criteria were used for diagnosing IIMs and classifying DM and PM [22]. Classification of clinically amyopathic DM (CADM) was based on the modified Sontheimer criteria [23]. The diagnosis of ILD was based on the American Thoracic Society criteria, which included consistent clinical features and pulmonary function tests, radiographic evidence of interstitial disease, and/or lung histopathology consistent with the diagnosis [24]. Finally, 80 patients with longitudinal HRCT scans were analysed; more than half of them (41 patients, 51.3%) had three longitudinal scans. The study was conducted according to the guidelines of the Declaration of Helsinki and the study design was approved by the Institutional Review Boards of Seoul National University Hospital (IRB No. 1902–091-1010) and Seoul National University Hospital Bundang Hospital (IRB No. B-1706/402–116). Informed consent was waived because of the retrospective nature of the study, and the analysis used anonymous clinical data.

Data collection

Clinical information at the time of the first HRCT scan (baseline) and at the time of IIMs diagnosis was collected through a medical chart review regarding demographic variables; clinical and laboratory data; the use of medications such as glucocorticoids and immunosuppressive or immunomodulatory drugs, including CYC, AZA, MMF, methotrexate, ciclosporin, tacrolimus, rituximab and IVIG; and lung transplant-free survival. Pulmonary function tests included the percentage of predicted forced vital capacity (FVC) (%FVC), and the percentage of predicted diffusing capacity for carbon monoxide (DLCO) (%DLCO). Laboratory variables included CRP, mg/dl levels, ESR, mm/h, lactate dehydrogenase (LDH, IU/l) levels, creatine phosphokinase (CPK, IU/l) levels and seropositivity for the anti-nuclear antibody and anti-Jo 1 antibody.

Visual pattern analyses of chest HRCT

Visual ILD patterns were analysed from baseline HRCT scans based on previous studies [25, 26] and initially assessed by two experienced thoracic radiologists (S.H.Y and L.P.) blinded to the patient's clinical information. Discordant readings (13 cases, 16%) were adjudicated by a third experienced thoracic radiologist (J.Y.K). The agreement between the two readings after adjudication was substantial ($\kappa = 0.74$, 95% CI = 0.61, 0.87) [27]. The assessment of ILD components—reticulation, ground-glass opacity and honeycombing—was scored using Likert scale semi-quantitative scores, ranging from 0 to 5 in the anatomical lung lobes. The intraclass correlation coefficient of these measurements was 0.81.

Automated quantification of chest HRCT analyses

Quantitative analysis of HRCT images was conducted by the Radiology Core at the University of California, Los Angeles, as described previously [14], and detailed information was summarized in Supplementary Method S1, available at *Rheumatology* online. Radiomic features analyses have been reported to be reproducible for the slice thickness ≤2 mm [28], and the thicknesses of all HRCT scans in our study were 1.25 mm or less. Regarding follow-up CT scans, half of the subjects had the same manufacturers and reconstruction kernels of scanners. The rest used similar sharp kernels despite different CT scanners, and we used normalized denoise technique prior to the calculation of radiomic features in order to overcome these differences. QILD scores were defined as the sum of quantitative lung fibrosis (QLF, %), quantitative ground-glass opacity (QGG, %) and quantitative honeycombing (QHC, %). The QILD scores of the whole lung (WL) and the most severe zone (MSZ) were used for the analysis. Because of the different time intervals between two consecutive HRCT scans, individual time-estimated changes in the QILD scores were derived using a linear approximation of yearly changes. Radiographic ILD progression was defined as changes in QILD scores over 2%, which was determined by performing statistical analysis using quantitative imaging biomarkers as in a previous study [29].

Statistical analyses

Continuous variables were summarized as means with SD. in a normal distribution or as medians with interquartile ranges (IQR) in a non-normal distribution. Categorical variables were summarized as absolute numbers and percentages. Data comparisons were conducted using the Student's t test or Mann–Whitney U test and χ^2 or Fisher's exact test. The relationship between pulmonary function parameters and QILD scores was tested using the Pearson correlation analysis. Kaplan–Meier survival curves with the log-rank test were used to analyse data related to lung transplant-free survival according to the two groups stratified median value of baseline WL-QILD, WL-QLF or WL-QGG scores. The follow-up period was calculated from the initial HRCT scan to the date of lung transplantation/death or censoring time (date of vital status ascertainment: 31 July 2021). The multivariable logistic regression model was used to find the independent related factor for radiographic ILD progression (ΔQILD >2%) among the variables that had statistical significances in univariable analyses or clinical implications. P-values of ≤0.05 were considered statistically significant. All statistical analyses and data visualization were performed using R software (version 4.1.1; R Foundation for Statistical Computing, Vienna, Austria) and Prism (version 8.0.1; GraphPad Software, San Diego, California, USA) for Windows 10.

Results

Characteristics of the included patients with IIMs-ILD

Eighty patients with IIMs-ILD who underwent serial HRCT scans at least twice were analysed. The mean (S.D.) age of the patients was 51.5 (11.7) years, and 60 (75.0%) were women. Of the 80 patients, 12 patients (15%) were diagnosed with ILD prior to IIMs diagnosis. The mean follow-up period was 7.1 (3.4) years. In most patients, the MSZ was the left or right lower zone. The baseline median (IQR) WL-QILD score and MSZ-QILD score were 28.1% (19.1–43.8) and 68.0% (45.5–81.8), respectively. In the visual pattern analysis, the most common ILD subtype was NSIP (57.5%), followed by organizing pneumonia (OP) (21.2%) and UIP (18.8%). The mean %FVC was 68.3 (17.3), and the mean %DLCO was 57.2 (18.7). Details of other clinical and laboratory characteristics are shown in Table 1. In the subgroup analysis based on the positivity to anti-Jo-1 antibodies, higher QILD scores in both WL and MSZ were observed in the anti-Jo-1 positive group ($n = 17$), but the differences were not statistically significant (Supplementary Table S1, available at *Rheumatology* online).

Correlations between QILD scores and pulmonary function test parameters in IIMs-ILD

The relationships between QILD scores and pulmonary function indices are presented in Fig. 1. The WL- and MSZ-QILD scores showed statistically significant correlations with %FVC ($r = -0.349$, $P = 0.002$ for WL-QILD; $r = -0.354$, $P = 0.002$ for MSZ-QILD) and %DLCO ($r = -0.381$, $P = 0.001$ for WL-QILD; $r = -0.340$, $P = 0.003$ for MSZ-QILD).

Individual QILD changes over two scans in IIMs-ILD

The median interval of two HRCT scans was 11.4 (4.8–19.6) months. We assessed the individual time-estimated yearly changes in the QILD scores between the first and second scans in the WL and MSZ (Fig. 2). Approximately half of the patients showed aggravation (increase) or stability (no changes) in WL-QILD scores; however, when patients were sorted by visual ILD subtype on HRCT, approximately two-thirds of the patients with UIP pattern had aggravated WL-QILD scores and less than half of the patients with NSIP and OP had aggravated scores (80% for UIP *vs* 44.4% for non-UIP, $P = 0.013$) (Fig. 2A). Similar patterns were observed for changes in MSZ-QILD scores (Fig. 2B). Next, the individual time-estimated yearly changes in the two major components of ILD-QGG and QLF are depicted in Fig. 2C (WL) and **2D** (MSZ). Interestingly, the changes of WL-QGG and WL-QLF tend to be subparallel, while those in the MSZ were mixed up. In particular, the deteriorated MSZ-QILD scores in the UIP pattern tended to be attributed to aggravated MSZ-QLF scores.

Radiographic ILD progression in IIMs-ILD

Among the 80 patients, 34 patients (42.5%) were categorised as a group with radiographic ILD progression (ΔQILD >2%). Table 2 compares the characteristics according to the radiographic ILD progression, showing that the UIP pattern ($P = 0.003$) and treatment with tacrolimus ($P = 0.043$) were frequently observed in the progression group. Subsequent multivariable logistic regression analysis identified that the UIP pattern was significantly associated with radiographic ILD progression (OR 6.67, 95% CI 1.57, 35.97, $P = 0.015$, Supplementary Table S2, available at *Rheumatology* online).

Table 1. Demographic and clinical characteristics of the patients with idiopathic inflammatory myopathies-associated interstitial lung disease (IIMs-ILD)

Characteristics	Total ($n = 80$)
Age at ILD diagnosis, years, mean (S.D.)	51.5 (11.7)
Female, *n* (%)	60 (75.0)
IIMs subtype, *n* (%)	
Polymyositis	22 (27.5)
Dermatomyositis[a]	58 (72.5)
Manifestations, *n* (%)	
Proximal muscle weakness, *n* (%) ($n = 79$)	55 (69.6)
Heliotrope rash ($n = 78$)	14 (17.9)
Gottron's papule ($n = 77$)	33 (42.9)
Raynaud phenomenon ($n = 77$)	18 (23.4)
Arthralgia/arthritis ($n = 78$)	40 (51.3)
Malignancies, *n* (%)	8 (10.0)
Laboratory findings at HRCT scan, median (IQR)	
CRP, mg/dL	0.3 (0.0–1.3)
ESR mm/h	30.0 (18.0–49.0)
CPK, IU/L	86.0 (39.0–760.0)
LDH, IU/L	293.0 (226.0–451.0)
Autoantibodies positivity, *n* (%)	
Anti-nuclear antibody	38 (48.7)
Anti-Jo-1 antibody	17 (22.1)
Pulmonary function, mean (S.D.)	
FVC, L ($n = 75$)	2.2 (0.7)
%FVC ($n = 75$)	68.3 (17.3)
DLCO, mL/min/mmHg ($n = 73$)	10.7 (4.0)
%DLCO ($n = 73$)	57.2 (18.7)
ILD subtype by HRCT, *n* (%)	
UIP	15 (18.8)
NSIP	46 (57.5)
OP	17 (21.3)
Others	2 (2.5)
Whole lung-QILD score, median (IQR)	
QILD, %	28.1 (19.1–43.8)
QGG, %	16.4 (9.7–23.0)
QLF, %	10.4 (5.8–17.3)
QHC, %	0.8 (0.2–1.8)
Most severe zone-QILD score, median (IQR)	
QILD, %	68.0 (45.5–81.8)
QGG, %	22.0 (14.4–30.1)
QLF, %	38.9 (19.5–56.5)
QHC, %	0.5 (0.1–1.5)
Most severe zone, *n* (%)	
Left lower zone	71 (88.8)
Right lower zone	8 (10.0)
Left middle zone	1 (1.2)
Acute exacerbation of ILD, *n* (%)	21 (26.3)
Lung transplantation, *n* (%)	2 (2.5)
Death, *n* (%)[b]	7 (8.8)

Data are presented as mean (S.D.), median (IQR) or *n* (%).
[a] Twelve (20.7%) patients with dermatomyositis were clinically amyopathic dermatomyositis.
[b] Causes of death included three acute exacerbation of ILD, two acute respiratory distress syndrome (one pneumonia; one diffuse alveolar haemorrhage), one septic shock after lung transplantation, and one refractory pulmonary arterial hypertension.
CPK: creatine phosphokinase; DLCO: diffusing capacity for carbon monoxide; FVC: forced vital capacity; HRCT: high-resolution CT; IIM: idiopathic inflammatory myositis; ILD: interstitial lung disease; IQR: interquartile range; LDH: lactate dehydrogenase; NSIP: non-specific interstitial pneumonia; OP: organizing pneumonia; QGG: quantitative ground-glass opacity; QHC: quantitative honeycombing; QILD: quantitative ILD score; QLF: quantitative lung fibrosis; UIP: usual interstitial pneumonia.

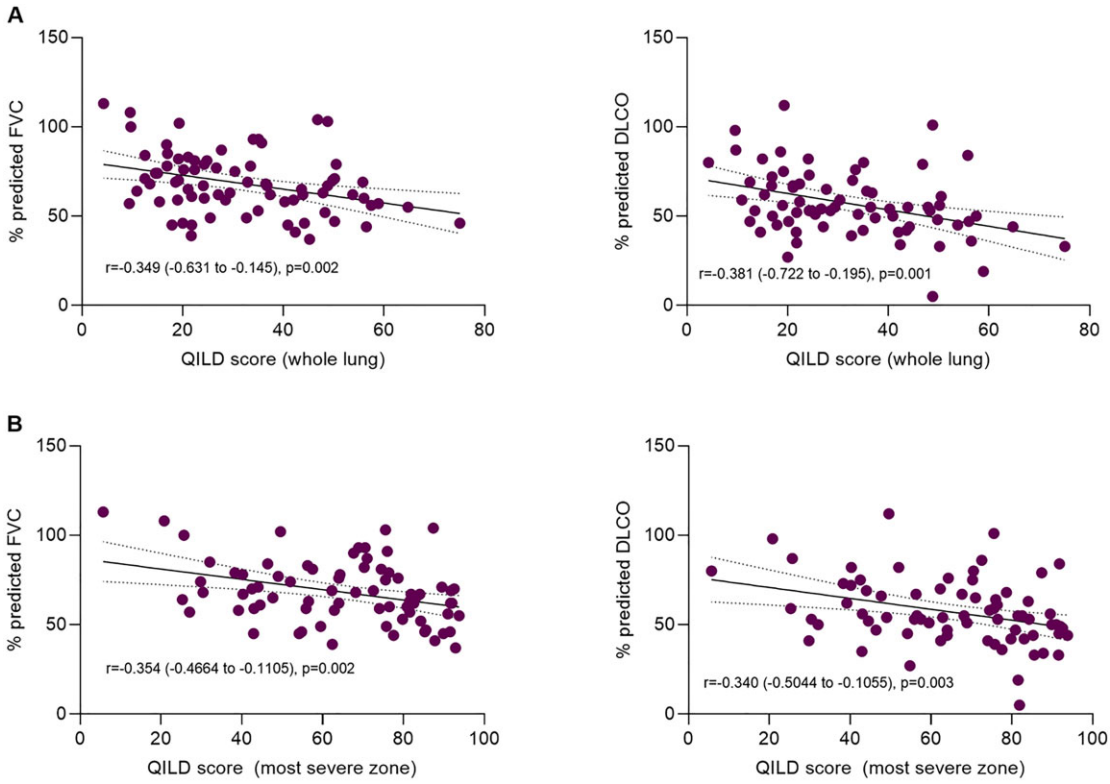


Figure 1. Correlations between QILD scores and pulmonary function indices at baseline. Scatter plots of QILD scores of the whole lung (**A**) and the most severe zone (**B**) with %predicted FVC (left panel) and %predicted DLCO (right panel). Pearson's correlation coefficients (r) with 95% CIs and *P*-values (*P*) are shown. DLCO: diffusing capacity of carbon monoxide; FVC: forced vital capacity; QILD: quantitative interstitial lung disease

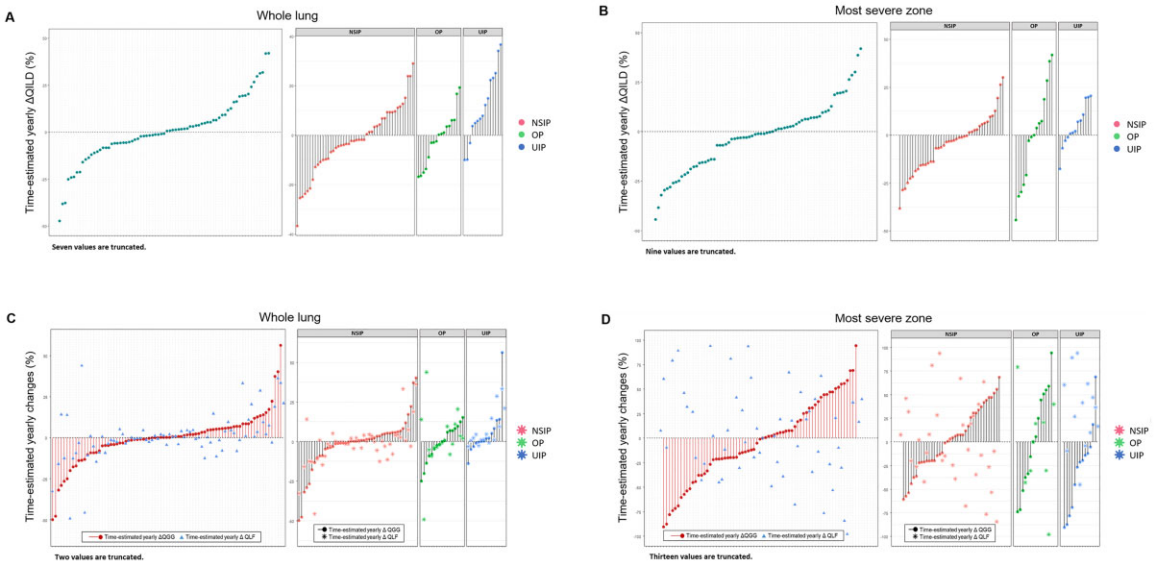


Figure 2. Individual time-estimated yearly ΔQILD scores between the first and second HRCT scans. Data are presented as Cleveland dot plots of ΔQILD scores in the whole lung (**A**) and most severe zone (**B**) and combined ΔQGG with ΔQLF scores (sorted by ΔQGG) in the whole lung (**C**) and most severe zone (**D**) for all patients (left panel) and in those sorted by ILD patterns (right panel). NSIP: non-specific interstitial pneumonia; OP: organizing pneumonia; QGG: quantitative ground-glass opacity; QILD: quantitative interstitial lung disease score; QLF: quantitative lung fibrosis; UIP: usual interstitial pneumonia

Association of baseline QILD scores and transplant-free survival in IIMs-ILD

Among the 80 patients, two (2.5%) underwent lung transplantation, and seven (8.8%) died during follow-up due to ILD-related complications. Causes of death included three acute exacerbations of ILD, two acute respiratory distress syndromes (one pneumonia and one diffuse alveolar haemorrhage), one septic shock after lung transplantation and one refractory pulmonary arterial hypertension. Transplant-free survival based on baseline QILD, QLF and QGG scores by Kaplan–Meier analyses are demonstrated in Supplementary Fig. S1, available at *Rheumatology* online. Patients with

Table 2. Comparison of characteristics according to the whole lung QILD progression (ΔQILD >2%)

	Non-progressor ($n = 46$)	Progressor ($n = 34$)	*P*-value
Age at ILD diagnosis, years, mean (S.D.)	52.8 (12.9)	49.9 (9.6)	0.278
Female, *n* (%)	36 (78.3)	24 (70.6)	0.601
IIMs subtype, *n* (%)			0.667
Polymyositis	14 (30.4)	8 (23.5)	
Dermatomyositis	32 (69.6)	26 (76.5)	
Malignancies, *n* (%)	8 (17.4)	0 (0.0)	N/A
Laboratory findings at baseline CT scan			
ESR, mm/h, median (IQR)	35.5 (20.0–51.5)	23.0 (17.0–42.0)	0.120
CRP, mg/dL, median (IQR)	0.2 (0.0–1.8)	0.3 (0.0–1.1)	1.000
CPK, IU/L, median (IQR)	84 (39–813)	101 (47–270)	0.779
LDH, IU/L, median (IQR)	295 (239–499)	291 (222–350)	0.346
Anti-Jo1 positivity, *n* (%)	11/45 (24.4)	6/32 (18.8)	0.753
Pulmonary function			
FVC, L, median (IQR)	2.1 (1.8–2.6)	2.1 (1.8–2.7)	0.858
%FVC <60, *n* (%)	11/45 (24.4)	13/30 (43.3)	0.143
DLCO, mL/min/mmHg, mean (S.D.)	10.3 (4.0)	11.4 (3.9)	0.247
%DLCO <40, *n* (%)	5/44 (11.4)	5/30 (16.7)	0.757
Visual ILD pattern—UIP	3 (6.5)	12 (35.3)	**0.003**
Overall extent of ILD, *n* (%)			0.854
<10%	11 (23.9)	7 (20.6)	
10–20%	5 (10.9)	4 (11.8)	
>20%	30 (65.2)	23 (67.6)	
Treatment between two scans, *n* (%)			
High-dose glucocorticoid	13 (28.3)	9 (26.5)	1.000
Ciclosporin A	19 (41.3)	9 (26.5)	0.255
Cyclophosphamide	3 (6.5)	4 (11.8)	0.674
Azathioprine	13 (28.3)	8 (23.5)	0.827
Tacrolimus	1 (2.2)	6 (17.6)	**0.043**
Mycophenolate mofetil	4 (8.7)	7 (20.6)	1.000
Methotrexate	4 (8.7)	7 (20.6)	0.231
IVIG	1 (2.2)	3 (8.8)	0.406
Rituximab	1 (2.2)	1 (2.9)	1.000

Data are presented as mean (S.D.), median (IQR) or *n* (%). Comparisons between two groups are conducted using the Student's *t test*, Mann–Whitney *U* test, χ^2 or Fisher's exact test. Bold values indicate significant *P*-values <0.05.
CPK: creatine phosphokinase; DLCO: diffusing capacity for carbon monoxide; FVC: forced vital capacity; IIM: idiopathic inflammatory myositis; ILD: interstitial lung disease; IQR: interquartile range; LDH: lactate dehydrogenase; QILD: quantitative ILD score.

higher baseline QILD (>28.1%) or QLF (>10.4%) scores had a higher risk of lung transplantation or death within 16 years ($P = 0.018$ for QILD and $P = 0.026$ for QLF, respectively).

Changes in QILD scores according to used medications

Changes in the QILD scores between the first and second HRCT scans according to the medications used are shown in Supplementary Table S3, available at *Rheumatology* online. There were no immunosuppressive drugs related to a significant improvement in the QILD scores, while we observed an aggravation of QILD scores in tacrolimus users ($n = 7$) compared with tacrolimus non-users ($n = 73$) (median changes in estimated WL-QILD +20.3 (2.7–38.4) in tacrolimus users *vs* −1.2 (−8.3–6.5) in tacrolimus non-users, $P = 0.013$, Supplementary Fig. S2, available at *Rheumatology* online).

Four distinct patterns in the three serial HRCT scans

Of the 80 patients, 41 underwent three serial HRCT scans. These patients were subdivided into four distinct groups based on the changing patterns of QILD scores in three consecutive scans: consistent improvement ($n = 7$) or worsening ($n = 7$), and two undulating courses: convex-like dynamic change ($n = 11$); and concave-like dynamic change ($n = 16$) (Fig. 3A). Time-estimated yearly changes in QILD scores between scans for each group were calculated and visualized in Fig. 3B. However, there were no differences between the four groups with respect to clinical characteristics (Supplementary Table S4, available at *Rheumatology* online).

Individual case reviews supported defined pattern changes. Fig. 4 shows a representative patient of three longitudinal scans that displayed a concave-like dynamic change, where the HRCT was performed at baseline, 6 months and 26 months. The patient was diagnosed with dermatomyositis and treated with glucocorticoids plus tacrolimus during subsequent follow-up. Lung volumes were stable over time, and the corresponding WL-QILD scores were 22.4% at baseline, 46.3% at 6 months and 35.5% at 26 months, respectively.

Discussion

In the present study, we evaluated QILD scores as a useful imaging biomarker for reflecting physiologic lung function, identifying changing patterns of serial scans, and predicting prognosis based on longitudinally obtained HRCT scans of IIMs-ILD. The QILD scores of IIMs-ILD showed significant negative correlations with pulmonary function parameters, including FVC and DLCO. The longitudinal changes in QILD scores were dynamic and differed according to the ILD visual subtypes. The UIP pattern was an independent predictive factor for radiographic ILD progression over two consecutive

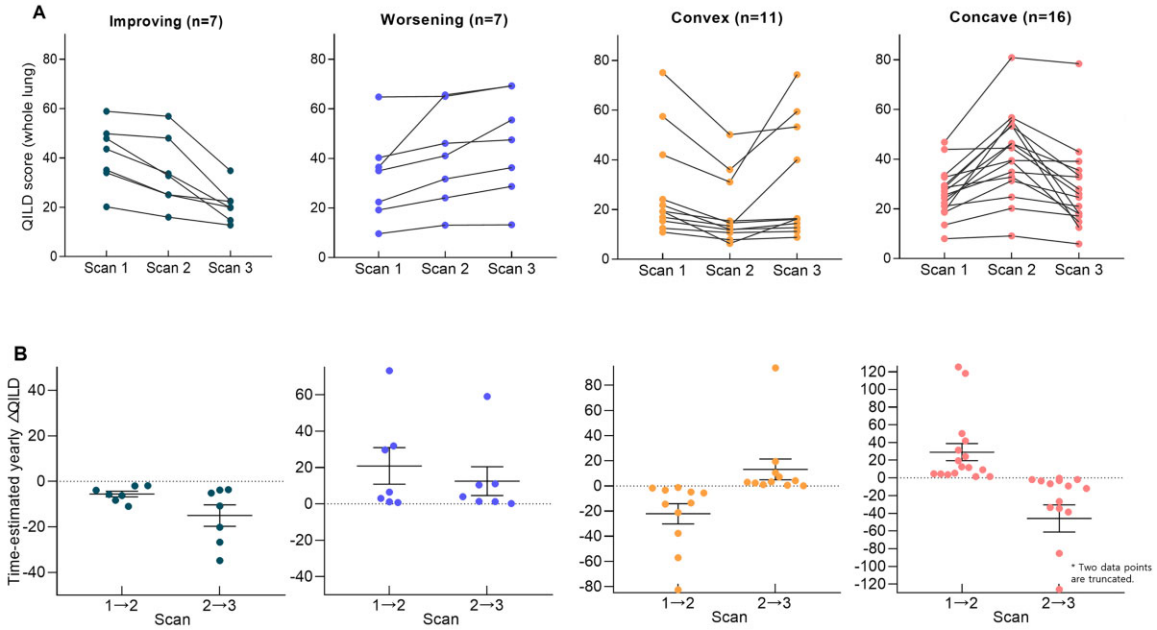


Figure 3. Four distinct patterns of dynamic ΔQILD scores from three consecutive HRCT scans ($n=41$). Whole lung-QILD scores (**A**) and yearly time-estimated ΔQILD scores between scans (**B**) are described as follows: improving ($n=7$), worsening ($n=7$), convex-like change ($n=11$), and concave-like change ($n=16$). HRCT: high-resolution CT; QILD: quantitative interstitial lung disease

scans. Certain baseline QILD scores may be associated with lung transplant-free survival; hence, they can help distinguish the poor prognostic group early.

Early detection and serial monitoring of treatment responses are crucial in managing IIMs-ILD. Pulmonary function tests are helpful in assessing respiratory impairment, disease progression and prognosis of ILD. Our study showed that QILD scores were significantly negatively correlated with %FVC and %DLCO (Fig. 1). It suggests that QILD scores are a useful outcome measure for evaluating serial HRCT in IIMs-ILD, as in RA-ILD or SSc-ILD [17, 30]. Because the involvement of respiratory muscle makes respiratory evaluations difficult when IIMs are severe, it may result in a modest correlation between QILD scores and lung function indices [6].

Quantitation of ILD parenchymal changes on HRCT images using automated software is an attractive tool for measuring the extent, volume and pattern of ILD and monitoring their changes over time. To date, several automated quantitation tools for assessing IIMs-ILD have been reported, including ours and the Computer-Aided Lung Informatics for Pathology Evaluation and Rating (CALIPER) [21, 31–33]. Ungprasert *et al.* analysed 110 patients with IIMs-ILD using the CALIPER method and showed good correlations between CALIPER measurements and functional parameters, similar to our results [33]. They performed CALIPER measurements on HRCT scans at baseline and 1 year; however, follow-up CT was performed only in 48 patients (43.6%) and did not evaluate their relevance to clinical outcomes.

In the present study, when individual changes in QILD scores were sorted by visual assessment of the ILD subtype on HRCT, more patients experienced an aggravation of QILD scores in the UIP group than in the non-UIP group, including NSIP and OP (80% for UIP *vs* 44.4% for non-UIP; Fig. 2A and B). This is consistent with a previous study showing that the UIP pattern was worse than that of NSIP or OP in disease progression [18]. In particular, the aggravated MSZ-QILD in the UIP group appeared to come from the increased MSZ-QLF score (Fig. 2D). This result is in line with a previous QILD study on SSc-ILD [34], showing the prominent transitional probability from QGG patterns to QLF patterns in the MSZ compared with the WL measurement. Although this QILD study did not include the transition of QILD components, future studies on the transitional probability of QILD measurement would provide further information on the pathobiology of IIM-ILD. In addition, another analysis of radiographic ILD progression (WL-ΔQILD >2%) confirmed the UIP pattern as an independent predictive factor (Table 2 and Supplementary Table S2, available at *Rheumatology* online).

Our study demonstrated that the baseline QILD scores in the WL and MSZ were 28.1% and 68.0%, respectively. We previously reported that the QILD scores in the WL and MSZ in SSc-ILD were 27.8% and 49.5%, respectively [16]. While the WL-QILD score was similar for both diseases, the MSZ-QILD score for IIMs-ILD was higher than that for SSc-ILD at baseline. Compared with QILD in RA-ILD [17], the baseline WL-QILD was similar in both diseases (WL-QILD in RA 26.7%), but dynamic changes in QILD scores were more pronounced in IIMs-ILD than in RA-ILD over three consecutive scans (Fig. 3).

In addition, higher baseline WL-QILD scores (>28.1%) and WL-QLF (>10.4%) were associated with a poor prognosis in the Kaplan–Meier curve with the log-rank test (Supplementary Fig. S1, available at *Rheumatology* online). Although our lung transplant-free survival analysis had low statistical power due to the small number of events for transplant or deaths, our study suggested the baseline QILD score as a potential prognostic marker for lung transplant-free survival. Further large-scale studies are warranted to evaluate the optimal cut-off of QILD scores and validate our findings.

In the absence of randomized clinical trials, treatments of IIMs-ILD are based on small retrospective studies [5, 6]. In

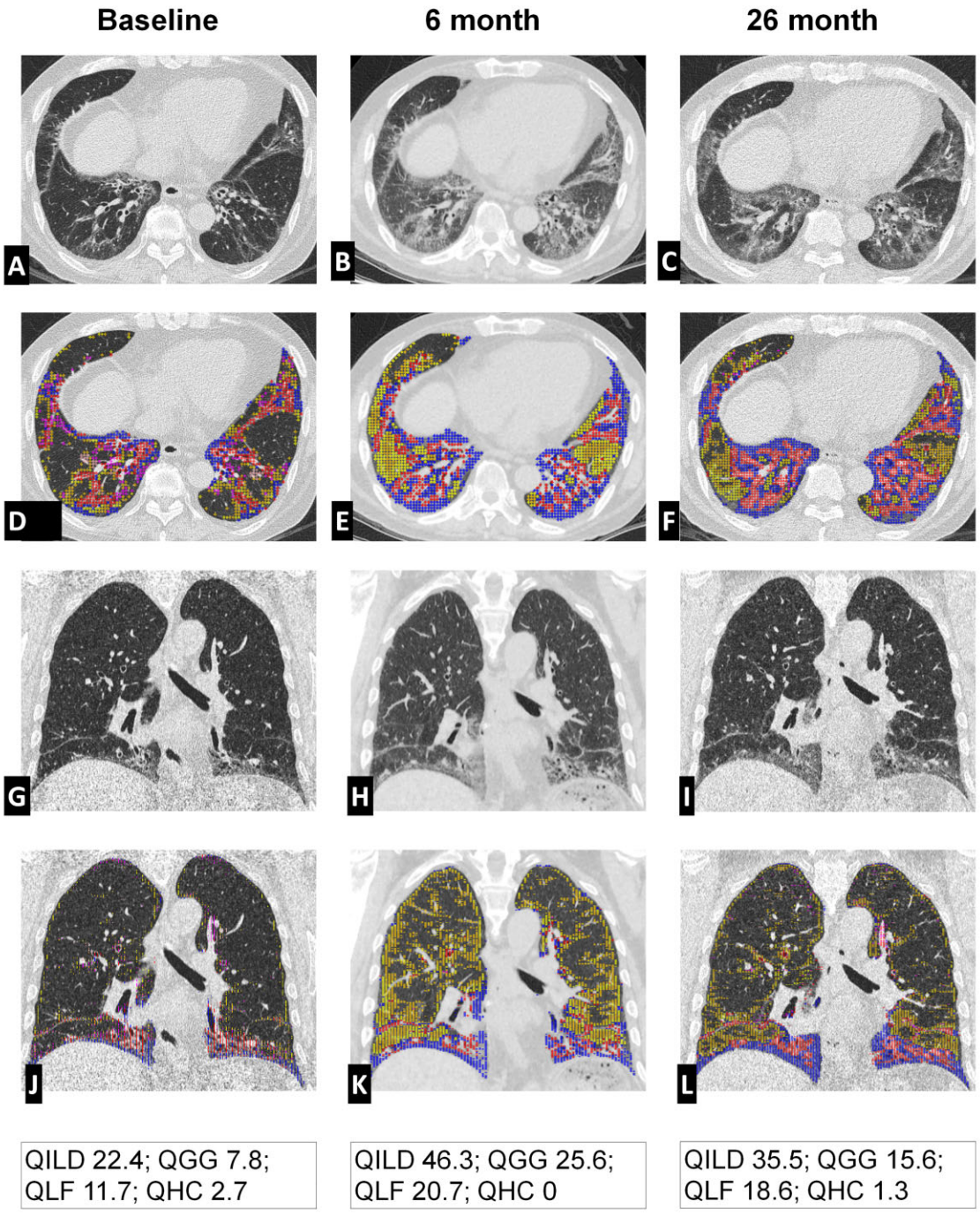


Figure 4. Representative serial images of a 57-year-old male patient with dermatomyositis showing concave-like dynamic QILD scores. Axial HRCT images are displayed consecutively at baseline, 6 months and 26 months (**A**–**C**). QILD scores are annotated on the axial images (**D**–**F**). Blue dots represent fibrosis, yellow dots represent ground-glass opacities and pink dots represent honeycombs. Coronal views of HRCT images (**G**–**I**) and QILD-annotated images (**J**–**L**) are also shown. Each QILD score is indicated in the box below. HRCT: high-resolution computed tomography; QGG: quantitative ground-glass opacity; QHC: quantitative honeycombing; QILD: quantitative interstitial lung disease score; QLF: quantitative lung fibrosis

the current study, we analysed longitudinal changes in QILD score according to the use of various immunosuppressive and immunomodulatory drugs, including CYC, AZA, MMF, methotrexate, ciclosporin, tacrolimus, IVIG, rituximab and high-dose steroids between baseline and the second HRCT scan. There were no immunosuppressive drugs associated with meaningful improvement in QILD scores during the period of the first and second HRCT scans. However, owing to the retrospective nature of our study, it is still difficult to determine the effect of certain drugs on QILD score changes based on our results. A representative patient treated with tacrolimus persistently and who underwent three longitudinal HRCT scans showed that WL-QILD scores showed concave-like changes over time (Fig. 4). To elucidate the relationship

between QILD changes according to these drugs, it is necessary to apply QILD scores to serial HRCT images from large, prospective and controlled studies.

This study had several limitations because of its retrospective nature in the study. First, because the HRCT scans were performed on demand by the clinician, the time intervals between the CT scans in each patient varied. Therefore, we used time-estimated changes in QILD scores between two consecutive scans using a linear approximation of yearly changes to adjust irregular follow-up intervals, but this could have led to over- or underestimation of QILD score changes. Second, various CT manufacturers or CT protocol changes were found in longitudinal CT scans over time in clinical follow-up. Consequently, different HRCT technical parameters naturally produced various noisy characteristics in HRCT images [28]. In this study, we applied an adaptive denoising technique with a standardized reference to removing various noise characteristics. Third, the relatively small number of patients from the two tertiary centres may have limited the generalisability of our results. Fourth, the events of lung transplantation or deaths were uncommon even after up to 16 years of follow-up. The thresholds of high-extent CT scores were chosen by baseline median QILD scores and applied in the same data set to explore the associations between clinical follow-up time and overall lung transplant-free survival. Further independent studies shall be followed to evaluate the optimal cut-off values of QILD scores and validity of estimated thresholds. Last, this study did not include information on myositis-specific antibodies such as anti-melanoma differentiation-associated protein 5 antibody, which is known as a marker for rapidly progressive ILD, because this test is not yet widely available in Korea. Nevertheless, considering the rarity of IIMs, our study included a considerable number of patients with IIMs-ILD. In addition, to the best of our knowledge, our study is the first to comprehensively investigate the longitudinal changes of IIMs-ILD in two or more serial HRCT scans using a quantitative ILD imaging tool.

In conclusion, the changes in QILD scores in IIMs-ILD are dynamic and differ according to the ILD subtype. The baseline visual UIP pattern was associated with QILD progression. Although the QILD system is not yet widely accessible, the QILD score may be a potential imaging biomarker for evaluating lung function and dynamic changes in the severity of ILD and prognosis in patients with IIMs-ILD.

Supplementary data

Supplementary data are available at *Rheumatology* online.

Data availability

The original contributions presented in the study are included in the article/Supplementary Material. The data underlying this study are available from the corresponding authors upon reasonable request.

Contribution statement

Y.-J.H. and G.-H.J.K. have full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. S.H.Y., J.S.L., E.Y.L., J.M.G., L.P., J.G.G., G.-H.J.K. and Y.-J.H. contributed to the study concept and design. J.Y., S.H.Y., J.Y.K., E.Y.L., J.G.G., J.M.G., L.P., G.-H.J.K. and Y.-J.H. contributed to the acquisition, analysis and interpretation of the data and contributed to the critical revision of the manuscript for important intellectual content. J.Y., G.-H.J.K. and Y.-J.H. drafted the manuscript and contributed to data interpretation. All authors have read and approved the final manuscript.

Funding

This work was supported by a grant awarded to Y.-J.H. from the SNUBH Research Fund (09–2017-004) and to G.-H.J.K. from NIH-NHLBI (R21HL 140465-01A1).

Disclosure statement: S.H.Y. works in MEDICALIP as a chief medical officer and has a stock in the firm outside this work. J.S.L. is a board member of GENOME INSIGHT, Inc. J.G.G. is a cofounder of MedQIA LLC. G.-H.J.K. is a research consultant to MedQIA LLC. All the other authors have no conflicts of interest to declare.

Acknowledgements

We would like to thank Editage (www.editage.co.kr) for English language editing.

References

1. Dalakas MC. Inflammatory muscle diseases. N Engl J Med 2015; 372:1734–47.
2. Fathi M, Dastmalchi M, Rasmussen E, Lundberg IE, Tornling G. Interstitial lung disease, a common manifestation of newly diagnosed polymyositis and dermatomyositis. Ann Rheum Dis 2004; 63:297–301.
3. Misra AK, Wong NL, Healey TT, Lally EV, Shea BS. Interstitial lung disease is a dominant feature in patients with circulating myositis-specific antibodies. BMC Pulm Med 2021;21:370.
4. Barba T, Fort R, Cottin V *et al.* Treatment of idiopathic inflammatory myositis associated interstitial lung disease: a systematic review and meta-analysis. Autoimmun Rev 2019;18:113–22.
5. Oddis CV, Aggarwal R. Treatment in myositis. Nat Rev Rheumatol 2018;14:279–89.
6. Hervier B, Uzunhan Y. Inflammatory myopathy-related interstitial lung disease: from pathophysiology to treatment. Front Med 2020; 6:326.
7. Wells AU, Desai SR, Rubens MB *et al.* Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. Am J Respir Crit Care Med 2003;167:962–9.
8. Bouros D, Wells AU, Nicholson AG *et al.* Histopathologic subsets of fibrosing alveolitis in patients with systemic sclerosis and their relationship to outcome. Am J Respir Crit Care Med 2002;165: 1581–6.
9. Tanizawa K, Handa T, Nakashima R *et al.* The prognostic value of HRCT in myositis-associated interstitial lung disease. Respir Med 2013;107:745–52.
10. Fujisawa T, Hozumi H, Kono M *et al.* Prognostic factors for myositis-associated interstitial lung disease. PLoS One 2014;9: e98824.
11. Kim HJ, Tashkin DP, Gjertson DW *et al.* Transitions to different patterns of interstitial lung disease in scleroderma with and without treatment. Ann Rheum Dis 2016;75:1367–71.
12. Laporte A, Mariampillai K, Allenbach Y *et al.* Idiopathic inflammatory myopathies: CT characteristics of interstitial lung disease and their association(s) with myositis-specific autoantibodies. Eur Radiol 2022;32:3480–9.

13. Kim HJ, Li G, Gjertson D *et al.* Classification of parenchymal abnormality in scleroderma lung using a novel approach to denoise images collected via a multicenter study. Acad Radiol 2008;15:1004–16.
14. Kim HG, Tashkin DP, Clements PJ *et al.* A computer-aided diagnosis system for quantitative scoring of extent of lung fibrosis in scleroderma patients. Clin Exp Rheumatol 2010;28(Suppl 62):S26–35.
15. Oh JH, Kim GHJ, Cross G *et al.* Automated quantification system predicts survival in rheumatoid arthritis-associated interstitial lung disease. Rheumatology 2022;61:4702–10.
16. Goldin JG, Kim GHJ, Tseng CH *et al.* Longitudinal changes in quantitative interstitial lung disease on computed tomography after immunosuppression in the Scleroderma Lung Study II. Ann Am Thorac Soc 2018;15:1286–95.
17. Lee JS, Kim GJ, Ha YJ *et al.* The extent and diverse trajectories of longitudinal changes in rheumatoid arthritis interstitial lung diseases using quantitative HRCT scores. J Clin Med 2021;10:3812.
18. Marie I, Josse S, Hatron PY *et al.* Interstitial lung disease in anti-Jo-1 patients with antisynthetase syndrome. Arthritis Care Res 2013; 65:800–8.
19. Aggarwal R, McBurney C, Schneider F *et al.* Myositis-associated usual interstitial pneumonia has a better survival than idiopathic pulmonary fibrosis. Rheumatology 2017;56:384–9.
20. Sugiyama Y, Yoshimi R, Tamura M *et al.* The predictive prognostic factors for polymyositis/dermatomyositis-associated interstitial lung disease. Arthritis Res Ther 2018;20:7.
21. Roncella C, Barsotti S, Valentini A *et al.* Evaluation of interstitial lung disease in idiopathic inflammatory myopathies through semi-quantitative and quantitative analysis of lung computed tomography. J Thorac Imaging 2022;37:344–51.
22. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). N Engl J Med 1975;292:344–7.
23. Sontheimer RD. Would a new name hasten the acceptance of amyopathic dermatomyositis (dermatomyositis siné myositis) as a distinctive subset within the idiopathic inflammatory dermatomyopathies spectrum of clinical illness? J Am Acad Dermatol 2002;46: 626–36.
24. Travis WD, Costabel U, Hansell DM *et al.* An official American Thoracic Society/European Respiratory Society statement: update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. Am J Respir Crit Care Med 2013;188: 733–48.
25. Ryerson CJ, Corte TJ, Lee JS *et al.* A standardized diagnostic ontology for fibrotic interstitial lung disease. an international working group perspective. Am J Respir Crit Care Med 2017;196:1249–54.
26. Goldin JG, Lynch DA, Strollo DC *et al.* High-resolution CT scan findings in patients with symptomatic scleroderma-related interstitial lung disease. Chest 2008;134:358–67.
27. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012;22:276–82.
28. Chong DY, Kim HJ, Lo P *et al.* Robustness-driven feature selection in classification of fibrotic interstitial lung disease patterns in computed tomography using 3D texture features. IEEE Trans Med Imaging 2016;35:144–57.
29. Obuchowski NA, Reeves AP, Huang EP *et al.* Quantitative imaging biomarkers: a review of statistical methods for computer algorithm comparisons. Stat Methods Med Res 2015;24:68–106.
30. Tashkin DP, Volkmann ER, Tseng CH *et al.* Relationship between quantitative radiographic assessments of interstitial lung disease and physiological and clinical features of systemic sclerosis. Ann Rheum Dis 2016;75:374–81.
31. Zavaletta VA, Bartholmai BJ, Robb RA. High resolution multidetector CT-aided tissue analysis and quantification of lung fibrosis. Acad Radiol 2007;14:772–87.
32. Marten K, Dicken V, Kneitz C *et al.* Interstitial lung disease associated with collagen vascular disorders: disease quantification using a computer-aided diagnosis tool. Eur Radiol 2009;19:324–32.
33. Ungprasert P, Wilton KM, Ernste FC *et al.* Novel Assessment of Interstitial Lung Disease Using the "Computer-Aided Lung Informatics for Pathology Evaluation and Rating" (CALIPER) Software System in Idiopathic Inflammatory Myopathies. Lung 2017;195:545–52.
34. Kim GHJ, Tashkin DP, Lo P *et al.* Using transitional changes on high-resolution computed tomography to monitor the impact of cyclophosphamide or mycophenolate mofetil on systemic sclerosis-related interstitial lung disease. Arthritis Rheumatol 2020;72: 316–25.