Biological correlates of radiological features of systemic sclerosis interstitial lung disease

Elizabeth R. Volkmann[1], Donald P. Tashkin[1], Mei Leng[1], Grace Hyun J. Kim ⓘ[2], Jonathan Goldin[2], Airi Harui[1] and Michael D. Roth[1]

[1]Department of Medicine, University of California, Los Angeles, David Geffen School of Medicine, Los Angeles, CA, USA. [2]Department of Radiology, University of California, Los Angeles, David Geffen School of Medicine, Los Angeles, CA, USA.

Corresponding author: Elizabeth R. Volkmann (evolkmann@mednet.ucla.edu)

Shareable abstract (@ERSpublications)
A comprehensive evaluation of bronchoalveolar lavage proteins suggests that radiological measures of fibrosis and ground-glass opacification may be linked to different underlying biological pathways in systemic sclerosis interstitial lung disease https://bit.ly/3ySClMF

Abstract

Background and objectives The extent and pattern of radiological features (*e.g.* fibrosis and ground glass) can influence treatment approaches for systemic sclerosis-related interstitial lung disease (SSc-ILD). However, the pathobiology underlying these radiological features is poorly understood and warrants further investigation.

Methods 68 proteins were measured in bronchoalveolar lavage (BAL) fluid from 103 SSc-ILD participants in Scleroderma Lung Study I. Quantitative image analysis calculated the extent of fibrosis (QLF) and ground-glass opacity (QGG) from concurrent high-resolution computed tomography (HRCT) scans. The relationship between BAL proteins and quantitative HRCT scores was assessed by univariate and multivariate analyses.

Results QLF scores correlated weakly with the extent of QGG, suggesting two distinct processes. In a univariate analysis, 25 proteins from several biological pathways correlated with QLF scores, including profibrotic factors, tissue remodelling proteins, proteins involved in monocyte/macrophage migration and activation, and proteins linked to inflammation and immune regulation. In contrast, QGG scores correlated with only six proteins, of which four were unique and related to granulocyte activation, mobilisation of bone marrow mesenchymal stem cells and activation of T-cells, B-cells, macrophages and eosinophils. In the multivariate models, interleukin-4, CCL7, receptor activator of nuclear factor-κB and tumour necrosis factor-α were independently associated with QLF, whereas interferon-γ was independently associated with QGG.

Interpretation QLF and QGG represent distinct radiological features of SSc-ILD, a conclusion reinforced by the presence of different biological pathways present within BAL fluid that associate with each. The identified proteins and related biological pathways may represent important therapeutic targets.

Introduction

High-resolution computed tomography (HRCT) imaging of the chest is considered the gold standard for diagnosing interstitial lung disease (ILD) in patients with systemic sclerosis (SSc) [1]. HRCT chest imaging also serves as a key monitoring tool to identify ILD progression [2]. Previous studies have demonstrated that the radiological extent of ILD predicts mortality in patients with SSc-ILD [3, 4]. Long-term follow-up of patients from the Scleroderma Lung Study (SLS) I and II also found that changes in the quantitative radiological extent of ILD over 1–2 years predicts long-term survival [5].

The predominant radiological features of SSc-ILD are reticulation with architectural distortion (fibrosis) and ground-glass opacity. While the presence of HRCT-defined fibrosis and ground-glass opacity are known to correspond with histological findings characterised by predominant fibrosis or inflammation,

respectively [6–8], the biological processes underlying these radiological features are largely unknown. However, in clinical practice, the presence of specific radiological features of ILD often guide treatment decisions with some practitioners favouring anti-inflammatory therapies when ground-glass opacity is the predominant pattern *versus* antifibrotic therapies when fibrosis is the predominant pattern.

Few studies have endeavoured to uncover the pathobiology behind distinct radiological features of SSc-ILD. Early research demonstrated that SSc-ILD patients with evidence of active alveolitis in bronchoalveolar lavage (BAL) fluid had more extensive ILD on HRCT of the chest [9–11]. Among patients enrolled in the multicentre SLS I [12], increased polymorphonuclear leukocytes and/or eosinophils in BAL fluid was associated with increased ground-glass opacity and fibrosis scores based on semi-quantitative analysis of concurrent HRCTs [10]. These findings are consistent with those of a single centre study, which demonstrated that patients with a higher eosinophil count in BAL fluid had increased extent of ground-glass opacity and fibrosis by semi-quantitative assessment [13]. Interestingly, this study also found that the concentration of certain BAL proteins correlated with the extent of ground-glass opacity (*e.g.* KL-6), whereas other BAL proteins correlated with the extent of fibrosis (*e.g.* CXCL8) [13]. The latter observation suggests that unique biological pathways may mediate the development of pulmonary interstitial inflammation *versus* fibrosis in SSc-ILD.

To further explore the biological pathways underlying distinct radiological features of SSc-ILD, the present study examined the correlations between the quantitative extent of ground-glass opacity (QGG) and fibrosis (QLF) and the concentration of 68 BAL proteins. We hypothesised that proteins known to mediate autoimmune response and inflammation would correlate more with the extent of QGG, whereas proteins involved in tissue remodelling and fibrosis would correlate more with the extent of QLF. The findings may reveal new therapeutic targets and provide further insight into the underlying biology of structural changes observed on HRCT in patients with SSc-ILD.

Material and methods

Study participants

Of the 158 participants who enrolled in SLS I [12] (NCT00004563), 148 underwent bronchoscopy with BAL during study screening [10]. SLS I included adult patients with SSc with alveolitis defined by any evidence of ground-glass opacity on HRCT and/or alveolitis on BAL fluid (defined as neutrophilia of $\geqslant 3\%$, eosinophilia of $\geqslant 2\%$ or both [14]) with a duration of disease $\leqslant 7$ years from onset of the first non-Raynaud phenomenon symptom of SSc [12]. Smoking within the previous 6 months was an exclusion criterion. Patients taking prednisone at a dose of greater than 10 mg per day, those who had previously been treated for more than 4 weeks with oral cyclophosphamide or who had received two or more intravenous doses, and those who had recently received other potentially disease-modifying medications were also excluded [10]. The Institutional Review Board of each site approved the studies, and only participants who provided informed consent were included in the present analyses.

Study design

Baseline measurements included pulmonary function tests, modified Rodnan Skin Score (mRSS) assessment, dyspnoea and health status evaluations and HRCT as previously described [12]. Quantitative image analysis was performed to calculate the extent of QLF and QGG at baseline in the whole lung (WL) and zone of maximum involvement (ZM) [15] (details are provided in supplementary material). In brief, QLF represents areas with a fibrotic reticulation pattern associated with architectural distortion of the lung. In contrast, QGG represents areas with pure ground-glass opacification with preserved underlying architectural features. As such, each pixel making up the computed tomography image can only have one label, representing the dominant feature assigned to that pixel. The scores for QLF and QGG are not overlapping, and higher scores indicate greater involvement by each feature in either the WL or ZM.

BAL protein analysis

Bronchoscopy with BAL was carried out in the right middle lobe (RML) in all subjects as previously described from 2000–2004 [10]. In previous studies, BAL performed in the RML was demonstrated as a suitable single-site for detecting alveolitis in SSc-ILD [16, 17], and single lavage performed in the RML determined the presence of alveolitis in 95% of patients with SSc-ILD [18]. Multiplex protein analyses were carried out for 68 different cytokines, chemokines, proteases and growth factors between 2007–2009 (additional details are in supplementary methods and supplementary table S1). We pre-selected proteins (based on available multiplex assays at that time) from different pathways implicated in SSc-ILD pathogenesis, aiming to include proteins involved in epithelial injury, angiogenesis, inflammation, autoimmunity and fibrosis.

Statistical analysis
Baseline characteristics
Summary statistics were generated for baseline patient characteristics. A two-sample t-test or Wilcoxon rank-sum test was used to compare continuous variables and a chi-squared test was used to compare categorical variables.

The distribution of each BAL protein was evaluated and corrected for skewness through logarithmic transformation where necessary. In the case where more than 80% of participants had unmeasurable values for a specific BAL protein (*i.e.* below the lower limit of detection), then that BAL protein was dichotomised as detectable *versus* undetectable.

Outcome measures
The primary outcome was the radiological extent of specific ILD features on HRCT (*i.e.* QGG-WL and QLF-WL). Pearson correlations were first performed to determine whether QGG-WL and QLF-WL scores were correlated with one another. Kendall's Tau correlations were then performed to examine the relationships between BAL protein levels and baseline QLF-WL and QGG-WL scores. Heatmaps were subsequently created to explore the interrelationships among proteins correlating with QLF-WL and with QGG-WL. Based on convention, any proteins correlating with QGG-WL/QLF-WL scores with $p<0.1$ were included in the heatmap analysis for the corresponding radiological feature.

Secondary outcomes included physiological measures of SSc-ILD severity, including forced vital capacity (FVC) % predicted, total lung capacity (TLC) % predicted and diffusion capacity for carbon monoxide (D_{LCO}) % predicted. Kendall's Tau correlations were performed to examine the relationship between BAL cell differentials (% eosinophils, neutrophils and lymphocytes) and QGG-WL and QLF-WL scores.

Multivariable models
Multivariable linear regression models were created to identify the key BAL proteins associated with QLF-WL and QGG-WL scores. Proteins that were found to correlate with QLF-WL and QGG-WL scores ($p<0.1$) on the univariable analysis were entered into the multivariable model using a backward selection process in combination with an examination of the Akaike information criterion (AIC) (see supplementary methods for additional details). The bootstrap procedure was employed for internal validation. Given the exploratory nature of this study, no corrections for multiple comparisons were performed. A threshold of $p<0.05$ was used to determine statistical significance in the multivariable models.

All tests were two-sided. All analyses were conducted in SAS version 9.4 (The SAS Institute, Cary, NC, USA).

Results
Participant characteristics
Among all SLS I participants, 103 participants had BAL specimens suitable for multiplex analysis. Compared with the entire SLS I cohort, participants who participated in the BAL analysis were similar in terms of demographic characteristics, as well as the severity of ILD (table 1).

Biomarkers correlating with radiological features of SSc-ILD
Relationship between QGG and QLF
The two principal HRCT features of SSc-ILD are fibrosis and ground glass. While we observed a statistically significant correlation ($p=0.028$) between QGG-WL and QLF-WL scores, the strength of their correlation was relatively weak ($r=0.19$) (figure 1).

Quantitative lung fibrosis
25 of the 68 measured proteins correlated ($p<0.1$) with QLF-WL or QLF-ZM scores in the univariable analysis (supplementary table S2 and figure 2). Proteins involved in tissue remodelling, including matrix metalloproteinases (MMP-1 ($r=0.13$), MMP-7 ($r=0.17$) and MMP-8 ($r=0.19$)) and hepatocyte growth factor (HGF; $r=0.29$) were positively correlated with QLF-WL scores. Proteins deemed as profibrotic mediators (*e.g.* transforming growth factor (TGF)-β; $r=0.14$ and platelet-derived growth factor (PDGF)-BB; $r=0.18$) were also positively correlated with QLF-WL scores. In addition, proteins involved in monocyte–macrophage migration and activation (chemokine (C-C motif) ligand 2 (CCL2; also known as MCP-1; $r=0.25$); CCL7 (also known as MCP-3; $r=0.21$); and macrophage colony-stimulating factor (M-CSF; $r=0.15$)) were positively correlated with QLF-WL scores. Of all of the proteins positively correlated with QLF scores, none overlapped with those that correlated with QGG scores. Figure 3 is a visualisation of the relationship between select proteins and QLF-WL scores, for which the correlation coefficient was >0.2.

	All SLS I participants (n=158)	SLS I participants in BAL analysis (n=103)
TABLE 1 Patient characteristics of the entire Scleroderma Lung Study (SLS) I cohort (n=158) and SLS I (n=103) participants who participated in this bronchoalveolar lavage (BAL) analysis		
Age, years	48.5±12.3	46.7±12.2
Female, %	70.3	73.8
Systemic sclerosis duration, years	2.7 (1.5–4.8)	2.9 (1.7–4.9)
Race, %		
White	63.9	62.1
African American	16.5	16.5
Asian	7.6	9.7
Other	12.0	11.7
mRSS	14.8±10.9	13.6±10.3
FVC % predicted	68.1±12.1	68.6±12.0
TLC % predicted	69.6±13.1	69.7±14.1
D_{LCO} **% predicted**	46.6±12.9	47.1±12.9
QGG % whole lung	23.4±9.5	23.4±9.7
QLF % whole lung	10.2±10.4	10.3±11.3
Alveolitis based on BAL parameters[#], %	63.9	57

Data are presented as mean±SD or median (interquartile range) unless otherwise stated. mRSS: modified Rodnan Skin Score; FVC: forced vital capacity; TLC: total lung capacity; D_{LCO}: diffusion capacity for carbon monoxide; QGG: quantitative extent of ground-glass opacity; QLF: quantitative extent of fibrosis. [#]: Alveolitis based on BAL parameters was defined as 3% or greater polymorphonuclear and/or 2% or greater eosinophilic leukocytes on lavage.

There were also proteins negatively correlated with QLF-WL or QLF-ZM scores (higher protein levels indicating less fibrosis). Interleukin (IL)-4 (r=−0.18), IL-12 (r=−0.18), IL-15 (r=−0.17), endoglin (r=−0.16), vascular endothelial growth factor (VEGF; r=−0.21), macrophage migration inhibitory factor
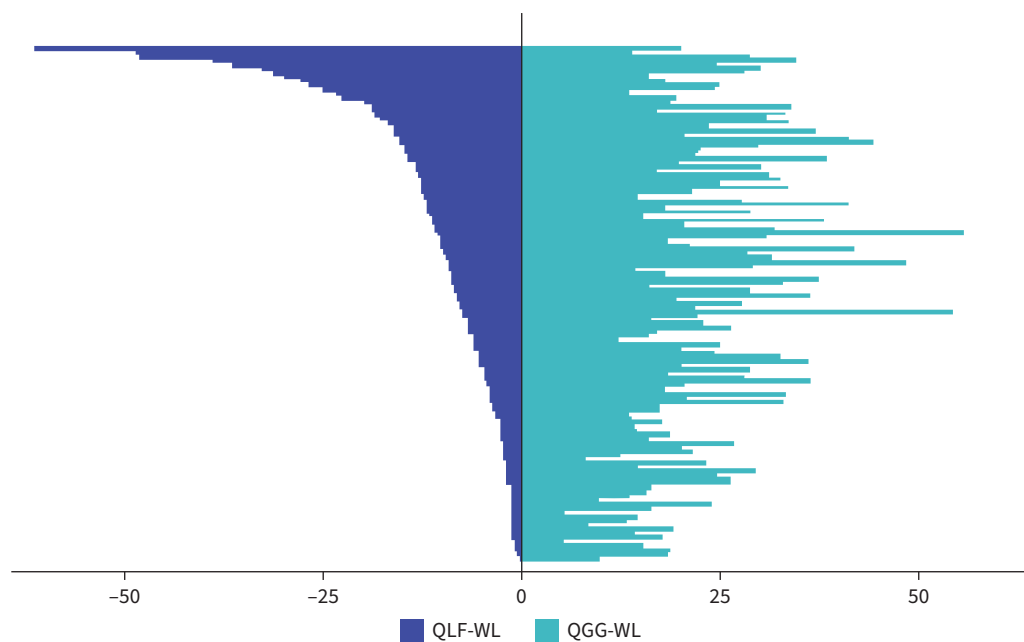

FIGURE 1 Relationship between QLF-WL scores (blue) and QGG-WL (green) scores for individual study participants. Each horizontal bar represents the QLF-WL and QGG-WL scores for an individual study subject. The Pearson correlation coefficient for QLF-WL and QGG-WL scores was 0.19, indicating a weak correlation (p=0.028). QGG: quantitative extent of ground-glass opacity; QLF: quantitative extent of fibrosis; WL: whole lung.
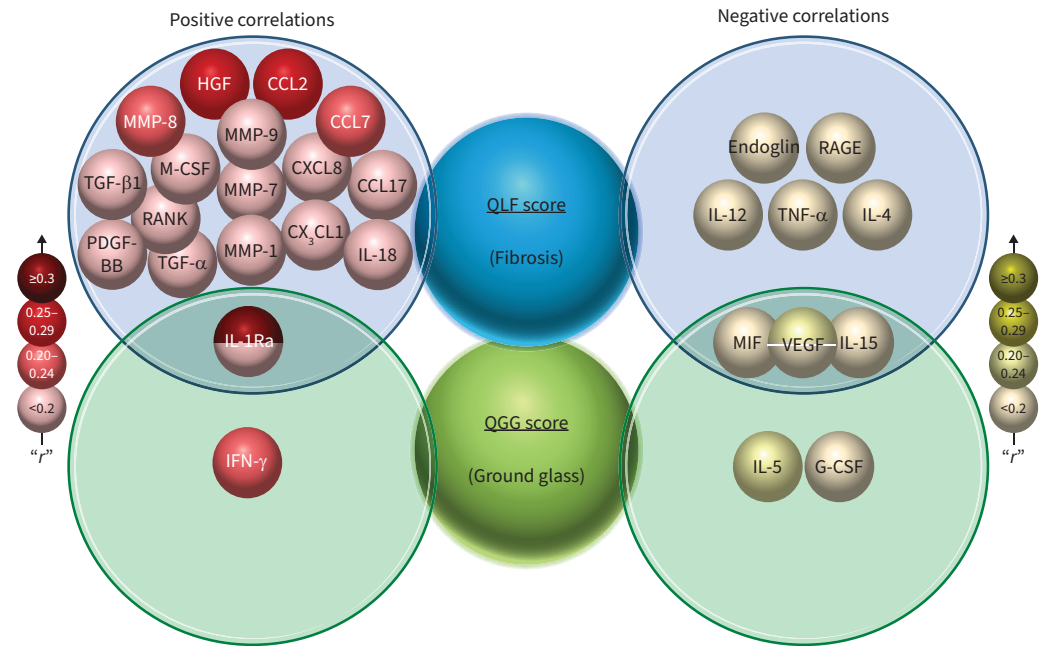
FIGURE 2 Bronchoalveolar lavage (BAL) proteins with positive (left side) and negative (right side) correlations with QLF-WL score (top) and QGG-WL (bottom). Darker shading around individual BAL proteins indicates a stronger correlation. Fewer BAL proteins were correlated with QGG-WL scores compared with QLF-WL scores. Only 4 among 68 measured BAL proteins correlated with both QGG-WL and QLF-WL scores. QGG: quantitative extent of ground-glass opacity; QLF: quantitative extent of fibrosis; WL: whole lung; TGF: transforming growth factor; MMP: matrix metalloproteinase; HGF: hepatocyte growth factor; PDGF: platelet-derived growth factor; IL: interleukin; RAGE: receptor for advanced glycation end products; G-CSF: granulocyte colony-stimulating factor; VEGF: vascular endothelial growth factor.

(MIF; r=−0.15), receptor for advanced glycation end products (RAGE; r=−0.14) and tumour necrosis factor (TNF)-α (r=−0.16) were negatively correlated with QLF-WL scores. VEGF (r=−0.15), IL-15 (r=−0.17) and MIF (r=−0.15) levels were also negatively correlated with QGG-WL scores (supplementary table S3). There was a high degree of concordance between proteins significantly correlated with QLF-WL scores and those with QLF-ZM scores (supplementary table S2).

Among the proteins correlating with QLF scores, a number of these proteins also correlated with pulmonary function test parameters, particularly D_{LCO} % predicted (supplementary table S4). MMP-1, MMP-7, MMP-8, CCL17 (also known as TARC), TGF-β1, PDGF-BB, IL-1Ra, chemokine (C-X-C motif) ligand 8 (CXCL8; also known as IL-8), CCL2, CCL7 and M-CSF were each negatively correlated with D_{LCO} % predicted, whereas VEGF, IL-15, IL-4 and IL-12 were each positively correlated with D_{LCO} % predicted.

When BAL proteins were assessed for their correlation to BAL cell differentials, the percentages of eosinophils (r=0.24), neutrophils (r=0.24) and macrophages (r=−0.20), but not lymphocytes, were significantly correlated with QLF-WL scores (p<0.05).

Ground-glass opacity
Six BAL proteins correlated (p<0.1) with QGG-WL or QGG-ZM scores in the univariable analysis (supplementary table S3 and figure 2). Interferon (IFN)-γ was positively correlated with QGG-WL (r=0.20) and QGG-ZM (r=0.23) scores, whereas IL-5 (r=−0.15), IL-15 (r=−0.17), granulocyte colony-stimulating factor (G-CSF; r=−0.17), MIF (r=−0.15) and VEGF (r=−0.15) were negatively correlated with QGG-WL. Figure 4 is a visualisation of the relationship between select proteins and QGG-WL scores, for which the correlation coefficient was ⩾0.15.

With respect to pulmonary function, only two of the six proteins were significantly correlated with FVC, D_{LCO} or TLC; IL-15 was positively associated with FVC % predicted (r=0.16), D_{LCO} % predicted (r=0.16)
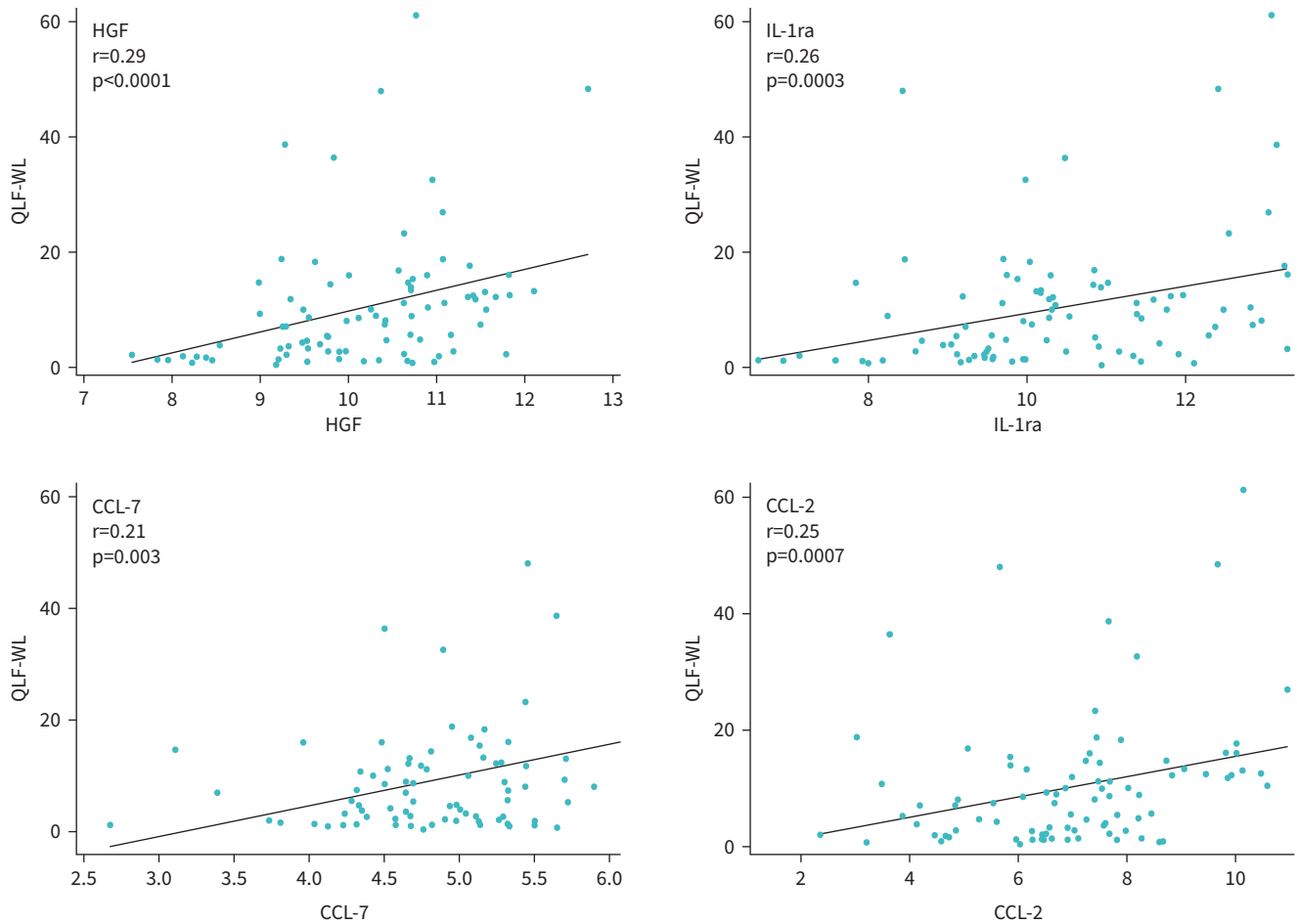
FIGURE 3 Visualisation of the relationship between select bronchoalveolar lavage (BAL) proteins and QLF-WL scores, for which the correlation coefficient was >0.2. For complete list of all BAL proteins significantly correlated with QLF-WL scores, please see supplementary table S2. QLF: quantitative image analysis for extent of fibrosis; WL: whole lung; HGF: hepatocyte growth factor; IL: interleukin; CCL: chemokine (C-C motif) ligand.

and TLC % predicted (r=0.18), whereas VEGF was positively associated with D_{LCO} % predicted (r=0.19) (supplementary table S4).

With respect to the BAL cell differential, the percentage of eosinophils was significantly correlated with QGG-WL (r=0.19; p=0.0124), but significant relationships were not observed between QGG-WL and the percentages of BAL neutrophils, lymphocytes or macrophages.

Heatmap analysis
A heatmap demonstrated intercorrelations among proteins that correlated with QLF-WL scores (figure 5a) and QGG-WL scores (figure 5b). For example, the QLF-WL heatmap depicts several distinct clusters of related mediators of fibrosis, including a cluster composed of M-CSF, MMP-8 and CXCL8; a large intercorrelated cluster, including MMP-1, receptor activator of NF-κB (RANK), TGF-β1, CCL17, CCL2 and HGF; and several smaller clusters composed of 2–3 intercorrelated proteins. The QGG-WL heatmap identifies a distinct cluster or intercorrelations formed by G-CSF, IL-15 and VEGF.

Multivariable analysis
The final multivariable model for QLF-WL score consisted of the following covariates: IL-4 presence, CCL7 level, RANK level, TNF-α level, HGF level, endoglin level and QGG-WL (table 2). Among these covariates, IL-4, CCL7, RANK and TNF-α remained significantly associated with QLF-WL score. QGG-WL was not significantly associated with QLF-WL scores in this final model. The adjusted R^2 for this model was 0.43 and the AIC was 472.99.
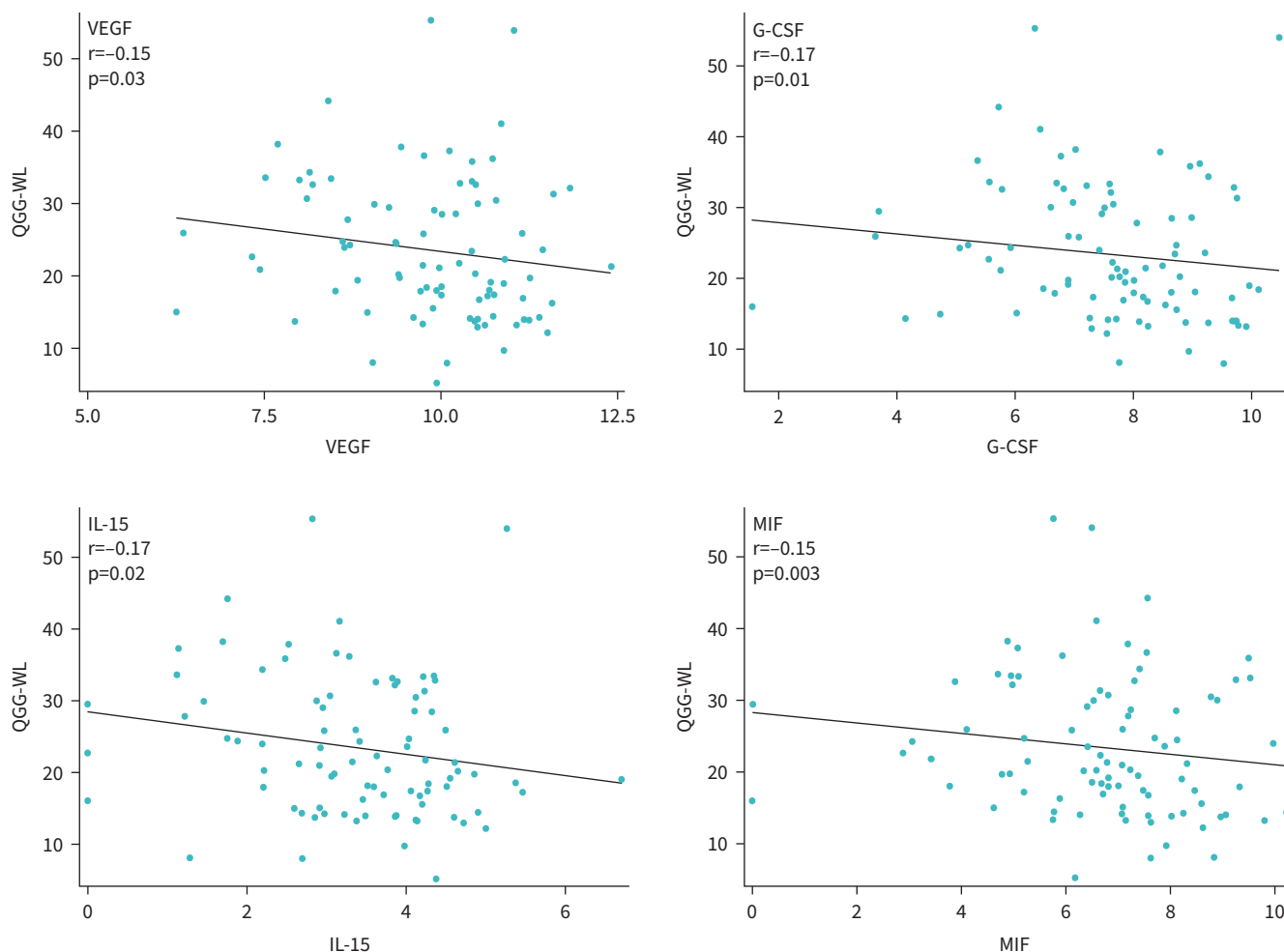
FIGURE 4 Visualisation of the relationship between select bronchoalveolar lavage (BAL) proteins and QGG-WL scores, for which the correlation coefficient was ⩾0.15. For complete list of all BAL proteins significantly correlated with QGG-WL scores, please see supplementary table S3. QGG: quantitative extent of ground-glass opacity; WL: whole lung; VEGF: vascular endothelial growth factor; G-CSF: granulocyte colony-stimulating factor; IL: interleukin; MIF: macrophage migration inhibitory factor.

A final multivariable correlation model for the QGG-WL included IL-5 presence, IL-15 level, IFN-γ presence and QLF-WL (table 3). Among these covariates, only IFN-γ remained significantly associated with QGG-WL score. QLF-WL was not significantly associated with QGG-WL scores in this final model. The adjusted R^2 for this model was considerably lower (0.16) and the AIC was 483.52.

Discussion

ILD affects the majority of patients with SSc and is a leading cause of death in patients with SSc [19, 20]. A deeper understanding of the underlying biology may lead to the discovery of new therapeutic targets and the opportunity to improve treatment outcomes [21]. The present study examined the relationship between the two main radiological features of SSc-ILD, fibrosis and ground-glass opacifications and then investigated how these features relate to biologically active proteins recovered from their lungs. First, we observed that QLF and QGG scores are only weakly correlated, suggesting that these features may arise from different biological processes. Second, we observed that each radiological feature correlated with its own distinct subset of proteins with limited overlap. Finally, this study identified interrelated networks of BAL proteins that might reflect biological pathways of key importance in the pathogenesis of SSc-ILD.

The first striking finding was that across the spectrum of SLS I participants, the extent of QGG was only weakly correlated with the extent of QLF. Moreover, in the multivariable models, no significant associations were observed between QLF and QGG scores. In support of this finding, a different repertoire of BAL proteins correlated with QLF scores than with QGG scores. Specifically, proteins involved in a
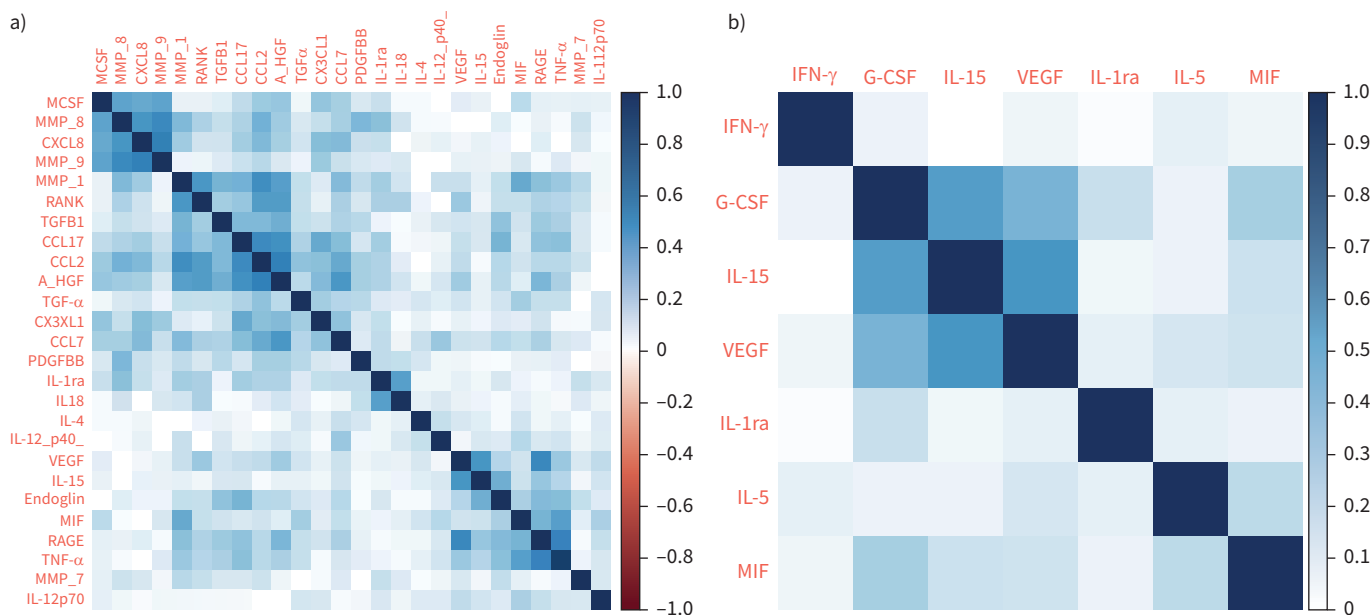
FIGURE 5 Intercorrelations among bronchoalveolar lavage (BAL) proteins correlating with a) QLF-WL scores and b) QGG-WL scores. Stronger intercorrelations are represented with darker shading. QGG: quantitative extent of ground-glass opacity; QLF: quantitative extent of fibrosis; WL: whole lung; M-CSF: macrophage colony-stimulating factor; MMP: matrix metalloproteinase; RANK: receptor activator of NF-κB; TGF: transforming growth factor; CCL: chemokine (C-C motif) ligand; HGF: hepatocyte growth factor; PDGF: platelet-derived growth factor; IL: interleukin; VEGF: vascular endothelial growth factor; MIF: macrophage migration inhibitory factor; RAGE: receptor for advanced glycation end products; TNF: tumour necrosis factor; IFN: interferon; G-CSF: granulocyte colony-stimulating factor.

TABLE 2 Bronchoalveolar lavage (BAL) proteins associated with QLF-WL in a multivariable linear regression model; the adjusted R^2 for this model was 0.43 and the AIC was 472.99

BAL protein	Estimate	95% CI	p-value
IL-4 presence	−4.86	−8.93– −0.79	0.020
CCL7 level	5.11	2.18–8.03	0.0008
RANK level	1.61	0.27–2.95	0.0192
TNF-α level	−2.59	−4.24– −0.94	0.0025
HGF level	2.02	−0.24–4.27	0.0794
Endoglin level	−0.01	−0.02–0.0	0.0896
QGG-WL	0.08	−0.13–0.28	0.460

AIC: Akaike information criterion; CI: confidence interval; IL: interleukin; CCL7: chemokine (C-C motif) ligand 7; RANK: receptor activator of NF-κB; TNF: tumour necrosis factor; HGF: hepatocyte growth factor; QGG: quantitative extent of ground-glass opacity; QLF: quantitative extent of fibrosis; WL: whole lung.

TABLE 3 Bronchoalveolar lavage (BAL) proteins associated QGG-WL in a multivariable linear regression model; the adjusted R^2 for this model was 0.16 and the AIC was 483.52

BAL protein	Estimate	95% CI	p-value
IL-15 level	−1.24	−2.83–0.35	0.126
IL-5 presence	−4.53	−9.18–0.12	0.0559
IFN-γ presence	4.30	0.29–8.30	0.0357
QLF-WL	0.15	−0.02–0.33	0.0827

QGG: quantitative extent of ground-glass opacity; WL: whole lung; AIC: Akaike information criterion; CI: confidence interval; IL: interleukin; IFN: interferon; QLF: quantitative extent of fibrosis.

number of biological pathways theorised to be associated with the pathogenesis of SSc-ILD (*e.g.* fibrosis, tissue remodelling, monocyte–macrophage migration and activation) correlated with QLF scores, suggesting that radiological fibrosis itself represents a dynamic process with diverse biological underpinnings. The finding that few proteins correlated with QGG scores was unexpected, but not entirely surprising. Of those factors that were either positively or negatively correlated with QGG scores, the majority were regulators of immunity, autoimmunity and inflammation. There is also evidence that QGG may not necessarily reflect the increased infiltration by inflammatory cells due to underlying SSc-ILD. For example, ground-glass opacification on HRCT may arise from architectural changes due to early fibrosis, as well as external factors, such as poor ventilation of dependent lung areas, oedema, airspace and interstitial pneumonia, and aspiration lung injury related to oesophageal dysmotility [22]. WELLS *et al.* [6] found that amorphous parenchymal opacification on HRCT was a less-reliable indicator of histological appearance than a reticular pattern. Another small study of patients with SSc-ILD found no relationship between the extent of ground-glass opacity and IL-6 and IL-7 levels in the BAL fluid of patients with SSc-ILD [23].

Known mediators of fibrosis, such as TGF-β and PDGF, were positively correlated with QLF scores, supporting the known pathobiology of SSc-ILD [24]. Specific mediators of matrix remodelling, including MMP-1, MMP-7, MMP-8 and MMP-9, were also positively correlated with QLF scores. These findings are consistent with peripheral proteomic profiles of patients with ILD. For example, previous studies have demonstrated increased genetic expression of MMP-7 [25] and increased levels of MMP-7 [26] in the circulation of patients with SSc-ILD. Moreover, higher peripheral blood levels of both MMP-7 and MMP-12 were observed in SSc patients with extensive ILD on HRCT [4] compared with limited ILD [26]. Collectively, these findings reaffirm the importance of matrix remodelling in SSc-ILD pathogenesis and also suggest that peripheral blood measurement of these proteins may serve as a surrogate for direct pulmonary measurement, although future studies are needed to test this hypothesis.

Of interest, the chemokines, CCL2 and CCL7 were also positively correlated with QLF-WL scores. Both chemokines are considered pro-inflammatory and contribute to tissue fibrosis by activating the synthesis of extracellular matrix proteins in fibroblasts. Observational studies have demonstrated that levels of peripheral CCL2 are higher in SSc patients with ILD [27, 28] and that increased CCL2 is associated with accelerated ILD progression [29]. Knockout animals missing the receptor for CCL2 are protected from developing severe pulmonary fibrosis after instillation of bleomycin [30], further supporting the role of CCL2 in ILD pathogenesis. While CCL7 is closely related to CCL2, CCL7 has chemotactic effects not only for monocytes, but also for dendritic cells and granulocytes [31]. Serum levels of CCL7 were elevated in SSc patients with ILD and associated with more severe restriction on pulmonary function testing [32].

Several BAL proteins were negatively correlated with QLF scores; among these proteins, VEGF was also negatively correlated with QGG scores. A key regulator of angiogenesis, VEGF plays a central role in lung development and maturation [33]. A previous analysis of BAL specimens from patients with SSc found that VEGF levels were lower in patients with alveolitis [34]. This study also demonstrated that VEGF levels were inversely correlated with QGG score on HRCT, consistent with the findings of the present study. There are a number of plausible explanations for these findings. For example, preclinical studies have demonstrated that VEGF-A protects against excessive pulmonary fibrosis resulting from lung injury [35, 36]. In addition, VEGF levels measured in the epithelial lining fluid of patients with acute respiratory distress syndrome (ARDS) were lower than in patients at risk for ARDS and inversely correlated with Lung Injury Score [37]. Studies have also demonstrated lower BAL fluid VEGF levels in patients with idiopathic pulmonary fibrosis [38]. However, treatments targeting VEGF (*e.g.* nintedanib) have been found to ameliorate pulmonary fibrosis in animal models [39] and slow the decline of lung function in human SSc-ILD studies [40]. Further research is needed to understand how VEGF and its unique isoforms contribute to the SSc-ILD pathogenesis.

The heatmap analysis for QLF-WL protein correlates demonstrated several clusters of interrelated proteins. For example, MMP-8, MMP-9, CXCL8 and G-CSF formed one cluster of related immune mediators of fibrosis. Intriguingly, the individual proteins significantly associated with QLF-WL scores in the multivariable analysis were each from different clusters on the heatmap. These findings suggest that understanding the relationships among different proteins is likely just as important as understanding the relationship between individual proteins and clinical outcomes. Had we solely presented the multivariable model, and not the heatmap analysis, one might erroneously deduce, for example, that VEGF and G-CSF are insignificant pathogenic mediators of radiological features of ILD.

The heatmap analysis for QGG-WL protein correlates also illuminated networks of intercorrelated proteins that may represent key pathogenic pathways. For instance, IL-15, G-CSF and VEGF (each individually associated with decreased QGG-WL scores) formed one cluster. Interestingly, in the multivariable analysis, IL-15 was significantly associated with QGG-WL score, whereas G-CSF and VEGF were not. This is likely because these three proteins were highly correlated with one another; the two other proteins included in the multivariable model were from other distinct clusters (IL-5 and IFN-γ). Future biomarker studies in this disease state should consider presenting heatmap analyses as they demonstrate important information about interrelated proteins that is not captured in multivariable modelling outputs.

The strengths of this study include our comprehensive evaluation of immune proteins from several biological pathways, in contrast with previous studies, which focused on a few select proteins measured in BAL fluid. Additional strengths include our use of a clinically well-characterised SSc-ILD cohort with a number of objective measures of ILD severity performed at the same time as bronchoscopy (e.g. radiological imaging and pulmonary function testing). Limitations include the fact that not all patients had suitable BAL specimens for multiplex analysis; however, reassuringly, the baseline characteristics of the included patients were similar to the entire study cohort. In addition, the current analysis was limited to measuring proteins that were available in multiplex formats at the time (2007–2009). Another important limitation was that corresponding blood specimens were not available to measure these proteins in the circulation, which may have provided an important comparison of the peripheral and pulmonary compartments and the opportunity to identify potential blood biomarkers. In addition, proteins in this study were all recovered from the RML and therefore, may not be entirely representative of disease activity occurring in other parts of the lungs. Given limitations in the HRCT protocol at the time, which captured axial slices at pre-defined anatomic zones, we were not able to directly compare BAL findings with ground-glass opacity and fibrosis scores in the RML. Moreover, without histopathological data, we cannot confirm the exact biological pathways underlying each radiological feature. Also, some of the HRCTs in SLS I were not volumetric, but axial skip, so sampling was not complete. Finally, as this study enrolled patients with early SSc-ILD with evidence of alveolitis, BAL protein analyses may not be representative of all patients with a diagnosis of SSc-ILD and further investigation may be needed in other SSc-ILD patient subgroups.

In summary, the present study demonstrated that specific immune pathways are associated with radiological features of ILD in SSc. Notably, more proteins were associated with QLF scores, suggesting that areas of reticulation and architectural distortion on HRCT represent a diverse array of biologically active processes. In addition to serving as important treatment targets, these proteins may also serve as predictors of progression of ILD and/or treatment response. Future studies are needed to determine whether measurements of these proteins in the circulation also correlate with distinct radiological features of ILD in patients with SSc.

This study is registered at www.clinicaltrials.gov with identifier number NCT00004563.

References

1   Rahaghi FF, Hsu VM, Kaner RJ, *et al.* Expert consensus on the management of systemic sclerosis-associated interstitial lung disease. *Respir Res* 2023; 24: 6.

2   Raghu G, Remy-Jardin M, Richeldi L, *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* 2022; 205: e18–e47.

3   Moore OU, Goh N, Corte T, *et al.* Extent of disease on high-resolution computed tomography lung is a predictor of decline and mortality in systemic sclerosis-related interstitial lung disease. *Rheumatology (Oxford)* 2013; 52: 155–160.

4   Goh NS, Desai SR, Veeraraghavan S, *et al.* Interstitial lung disease in systemic sclerosis: a simple staging system. *Am J Respir Crit Care Med* 2008; 177: 1248–1254.

5   Volkmann ER, Tashkin DP, Roth MD, *et al.* Early radiographic progression of scleroderma: lung disease predicts long-term mortality. *Chest* 2022; 161: 1310–1319.

6   Wells AU, Hansell DM, Corrin B, *et al.* High resolution computed tomography as a predictor of lung histology in systemic sclerosis. *Thorax* 1992; 47: 738–742.

7   Sumikawa H, Johkoh T, Fujimoto K, *et al.* Pathologically proved nonspecific interstitial pneumonia: CT pattern analysis as compared with usual interstitial pneumonia CT pattern. *Radiology* 2014; 272: 549–556.

8   Chung JH, Cox CW, Montner SM, *et al.* CT features of the usual interstitial pneumonia pattern: differentiating connective tissue disease-associated interstitial lung disease from idiopathic pulmonary fibrosis. *AJR Am J Roentgenol* 2018; 210: 307–313.

9   Volpinari S, La Corte R, Bighi S, *et al.* Bronchoalveolar lavage in systemic sclerosis with lung involvement: role and correlations with functional, radiological and scintigraphic parameters. *Rheumatol Int* 2011; 31: 1183–1188.

10   Strange C, Bolster MB, Roth MD, *et al.* Bronchoalveolar lavage and response to cyclophosphamide in scleroderma interstitial lung disease. *Am J Respir Crit Care Med* 2008; 177: 91–98.

11   Wells AU, Hansell DM, Rubens MB, *et al.* Fibrosing alveolitis in systemic sclerosis: bronchoalveolar lavage findings in relation to computed tomographic appearance. *Am J Respir Crit Care Med* 1994; 150: 462–468.

12   Tashkin DP, Elashoff R, Clements PJ, *et al.* Cyclophosphamide versus placebo in scleroderma lung disease. *N Engl J Med* 2006; 354: 2655–2666.

13   Hesselstrand R, Wildt M, Bozovic G, *et al.* Biomarkers from bronchoalveolar lavage fluid in systemic sclerosis patients with interstitial lung disease relate to severity of lung fibrosis. *Respir Med* 2013; 107: 1079–1086.

14   White B, Moore WC, Wigley FM, *et al.* Cyclophosphamide is associated with pulmonary function and survival benefit in patients with scleroderma and alveolitis. *Ann Intern Med* 2000; 132: 947–954.

15   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: S26–S35.

16   Miller KS, Smith EA, Kinsella M, *et al.* Lung disease associated with progressive systemic sclerosis. Assessment of interlobar variation by bronchoalveolar lavage and comparison with noninvasive evaluation of disease activity. *Am Rev Respir Dis* 1990; 141: 301–306.

17   Clements PJ, Goldin JG, Kleerup EC, *et al.* Regional differences in bronchoalveolar lavage and thoracic high-resolution computed tomography results in dyspneic patients with systemic sclerosis. *Arthritis Rheum* 2004; 50: 1909–1917.

18 Frigieri L, Mormile F, Grilli N, *et al.* Bilateral bronchoalveolar lavage in progressive systemic sclerosis: interlobar variability, lymphocyte subpopulations, and functional correlations. *Respiration* 1991; 58: 132–140.

19 Tyndall AJ, Bannert B, Vonk M, *et al.* Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. *Ann Rheum Dis* 2010; 69: 1809–1815.

20 Elhai M, Meune C, Avouac J, *et al.* Trends in mortality in patients with systemic sclerosis over 40 years: a systematic review and meta-analysis of cohort studies. *Rheumatology (Oxford)* 2012; 51: 1017–1026.

21 Volkmann ER, Andréasson K, Smith V. Systemic sclerosis. *Lancet* 2023; 401: 304–318.

22 Diederich S. Hochauflösende computertomographie der lunge: milchglas und seine differenzialdiagnosen [High resolution computed tomography of the lungs: ground glass opacity and its differential diagnosis]. *Radiologe* 2010; 50: 1141–1152.

23 Mittal S, Kumar U, Guleria R, *et al.* Bronchoalveolar lavage fluid cytokines in the assessment of interstitial lung disease in systemic sclerosis: a prospective study. *Eur Respir J* 2019; 54: Suppl. 63, PA1370.

24 Zhang Y, Distler JH. Therapeutic molecular targets of SSc-ILD. *J Scleroderma Relat Disord* 2020; 5: 17–30.

25 Xu Z, Chen W, Chen C. Matrix metalloproteinase 7 is a candidate biomarker in systemic sclerosis-associated interstitial lung disease. *Acta Reumatol Port* 2020; 45: 191–200.

26 Stock C, De Lauretis AD, Visca D, *et al.* Serum biomarkers in SSc-ILD: association with presence, severity and prognosis. *Eur Respir J* 2019; 54: Suppl. 63, PA5198.

27 Carulli MT, Handler C, Coghlan JG, *et al.* Can CCL2 serum levels be used in risk stratification or to monitor treatment response in systemic sclerosis? *Ann Rheum Dis* 2008; 67: 105–109.

28 Scala E, Pallotta S, Frezzolini A, *et al.* Cytokine and chemokine levels in systemic sclerosis: relationship with cutaneous and internal organ involvement. *Clin Exp Immunol* 2004; 138: 540–546.

29 Wu M, Baron M, Pedroza C, *et al.* CCL2 in the circulation predicts long-term progression of interstitial lung disease in patients with early systemic sclerosis: data from two independent cohorts. *Arthritis Rheumatol* 2017; 69: 1871–1878.

30 Okuma T, Terasaki Y, Kaikita K, *et al.* C-C chemokine receptor 2 (CCR2) deficiency improves bleomycin-induced pulmonary fibrosis by attenuation of both macrophage infiltration and production of macrophage-derived matrix metalloproteinases. *J Pathol* 2004; 204: 594–604.

31 Proost P, Wuyts A, Van Damme J. Human monocyte chemotactic proteins-2 and -3: structural and functional comparison with MCP-1. *J Leukoc Biol* 1996; 59: 67–74.

32 Yanaba K, Komura K, Kodera M, *et al.* Serum levels of monocyte chemotactic protein-3/CCL7 are raised in patients with systemic sclerosis: association with extent of skin sclerosis and severity of pulmonary fibrosis. *Ann Rheum Dis* 2006; 65: 124–126.

33 Barratt SL, Flower VA, Pauling JD, *et al.* VEGF (vascular endothelial growth factor) and fibrotic lung disease. *Int J Mol Sci* 2018; 19: 1269.

34 De Santis M, Bosello SL, Capoluongo E, *et al.* A vascular endothelial growth factor deficiency characterises scleroderma lung disease. *Ann Rheum Dis* 2012; 71: 1461–1465.

35 Kearns MT, Dalal S, Horstmann SA, *et al.* Vascular endothelial growth factor enhances macrophage clearance of apoptotic cells. *Am J Physiol Lung Cell Mol Physiol* 2012; 302: L711–L718.

36 Stockmann C, Kerdiles Y, Nomaksteinsky M, *et al.* Loss of myeloid cell-derived vascular endothelial growth factor accelerates fibrosis. *Proc Natl Acad Sci USA* 2010; 107: 4329–4334.

37 Thickett DR, Armstrong L, Millar AB. A role for vascular endothelial growth factor in acute and resolving lung injury. *Am J Respir Crit Care Med* 2002; 166: 1332–1337.

38 Meyer KC, Cardoni A, Xiang ZZ. Vascular endothelial growth factor in bronchoalveolar lavage from normal subjects and patients with diffuse parenchymal lung disease. *J Lab Clin Med* 2000; 135: 332–338.

39 Huang J, Maier C, Zhang Y, *et al.* Nintedanib inhibits macrophage activation and ameliorates vascular and fibrotic manifestations in the Fra2 mouse model of systemic sclerosis. *Ann Rheum Dis* 2017; 76: 1941–1948.

40 Distler O, Highland KB, Gahlemann M, *et al.* Nintedanib for systemic sclerosis associated interstitial lung disease. *N Engl J Med* 2019; 380: 2518–2528.