Journal Pre-proof

Early Radiographic Progression of Scleroderma Lung Disease Predicts Long-term Mortality

Elizabeth R. Volkmann, MD, Donald P. Tashkin, MD, Michael D. Roth, MD, Jonathan Goldin, MD, PhD, Grace H.J. Kim, MD

PII: S0012-3692(21)04435-4

DOI: https://doi.org/10.1016/j.chest.2021.11.033

Reference: CHEST 4740

To appear in: *CHEST*

Received Date: 3 September 2021

Revised Date: 12 November 2021

Accepted Date: 27 November 2021

Please cite this article as: Volkmann ER, Tashkin DP, Roth MD, Goldin J, Kim GHJ, Early Radiographic Progression of Scleroderma Lung Disease Predicts Long-term Mortality *CHEST* (2022), doi: https://doi.org/10.1016/j.chest.2021.11.033.

Text word count: 3482

Abstract word count: 291

Early Radiographic Progression of Scleroderma Lung Disease Predicts Long-term Mortality

Running head: Radiographic progression of SSc-ILD predicts mortality

Elizabeth R. Volkmann, MD[1]

Donald P. Tashkin, MD[1]

Michael D. Roth, MD[1]

Jonathan Goldin, MD, PhD[2]

Grace H.J. Kim, MD[2]

[1] Department of Medicine, University of California, Los Angeles, David Geffen School of Medicine

[2] Department of Radiology, University of California, Los Angeles, David Geffen School of Medicine

Corresponding author:

Elizabeth R. Volkmann, MD, MS

1000 Veteran Ave. Ste 32-59

Los Angeles, CA 90095

Email: evolkmann@mednet.ucla.edu

Conflicts of interest:

Elizabeth R. Volkmann: Consulting and speaker fees from Boehringer Ingelheim; Grant support from Kadmon, Horizon Therapeutics

Donald P. Tashkin: Grant support from Genentech/Roche

Michael D. Roth: Grant support from Genentech/Roche

Jonathan Goldin: None

Grace H.J. Kim: None

Funding information: NIH/NHLBI: K23 HL150237 (ERV), R01 HL089758 (DPT), R01 HL089901 (RME), U01 HL 60587 (DPT) and U01 HL 60606 (RME). Bristol—Myers Squibb supplied cyclophosphamide for use in Scleroderma Lung Study I and Hoffmann-LaRoche supplied mycophenolate mofetil for use in Scleroderma Lung Study II.

Notation of prior abstract/publication: Oral presentation at the virtual EULAR Congress, June 2021.

Key words: Systemic sclerosis, Interstitial lung disease, Mortality, Biomarkers

Abbreviation list

AIC:	Akaike information criterion
CYC:	Cyclophosphamide
DLCO:	Diffusing capacity for carbon monoxide
FVC:	Forced vital capacity
HRCT:	High-resolution computed tomography
ILD:	Interstitial lung disease
MMF:	Mycophenolate
mRSS:	Modified Rodnan skin score
QILD:	Quantitative interstitial lung disease
QLF:	Quantitative lung fibrosis
RCT:	Randomized controlled trial
SLS:	Scleroderma lung study
SSc:	Systemic sclerosis

Abstract

Background: Radiographic endpoints are commonly included in therapeutic trials for systemic sclerosis-interstitial lung disease (SSc-ILD); however, the relationship between these outcomes and long-term mortality is unclear.

Research Question: Do short-term changes in radiographic measures of ILD predict long-term survival in patients with SSc?

Study Design and Methods: Scleroderma Lung Studies (SLS) I and II evaluated the safety and efficacy of cyclophosphamide (SLS I and II) and mycophenolate (SLS II) for the treatment of SSc-ILD. Changes in the extent of ILD over time were assessed on high-resolution computed tomography scans of the chest by quantitative image analysis, an approach which applies a computer-based algorithm to objectively assess changes in the radiographic extent of ILD. Participants were subsequently followed for up to 12 (SLS I) and 8 (SLS II) years. Cox proportional hazards models determined whether the change in the quantitative radiographic extent of ILD predicted survival, adjusting for other known predictors of survival.

Results: Among SLS I and II participants, 82 and 90 had follow up imaging scans, respectively, and were included in the analysis. Participants in both trials who had an increase in their total quantitative radiographic extent of ILD scores of ≥2% at 12 months (SLS I) or 24 months (SLS II) had significantly worse long-term survival than those with change scores of <2% ($P \leq 0.01$, log rank test). In the multivariable Cox models, radiographic progression remained associated with worse long-term survival in SLS I (P=0.089) and II (P=0.014).

Interpretation: Data from two independent clinical trial cohorts with extensive long-term follow up demonstrate that radiographic progression of ILD over 12-24 months, in treatment and

placebo arms, can predict increased risk for long-term mortality in patients with SSc. These findings suggest that radiographic endpoints may serve as surrogates for mortality in SSc-ILD.

Interstitial lung disease (ILD) is the leading cause of death in systemic sclerosis (SSc)[1,2] and a major focus for therapeutic drug discovery in this field. Despite the burgeoning SSc-ILD therapeutic pipeline, there are no universally accepted endpoints that define an optimal treatment response. Endpoints that directly measure how a patient with SSc-ILD feels and functions are lacking, and while mortality is an unequivocal endpoint, most SSc clinical trials are inadequately powered and designed to use mortality as an endpoint.

The most commonly employed surrogate endpoint in SSc-ILD clinical trials is the forced vital capacity (FVC).[3-5] While studies have demonstrated that the course of FVC is related to survival in SSc-ILD,[6,7] the reliability of this parameter as a direct measure of lung disease may be limited in SSc, where extra-pulmonary manifestations can substantially affect its measurement and interpretation (e.g., cutaneous sclerosis involving the chest wall, respiratory muscle weakness).[8] Variations in pulmonary function test protocols, along with patient and technician effort, can further influence the reproducibility of the FVC in clinical practice and research. In addition, studies have demonstrated that the FVC correlates poorly with the actual radiographic extent of ILD in SSc.[9] Thus, there is an unmet need for the discovery of novel SSc-ILD study endpoints that are reliable, reproducible and can consistently predict mortality outcomes in SSc patients.

Objectively quantifying the radiographic extent of ILD may represent a more direct assessment of parenchymal lung disease burden and potentially obviate the impact of extra-pulmonary disease, patient effort and technical factors that can alter the measurement of FVC, as well as the intra- and inter-reader variability in visual radiographic change assessment.[10] Studies have demonstrated that increased radiographic extent of ILD at baseline predicts responsiveness to immunosuppressive therapy.[11,12] Additional studies have found that the radiographic ILD

endpoint is sensitive to change in SSc-ILD patients undergoing treatment.[13-16] However, no studies have evaluated whether a *change* in the radiographic extent of ILD predicts mortality in this population. To address this issue, the present study examined whether radiographic progression of ILD in SSc patients receiving treatment would predict long-term mortality. Using data from the Scleroderma Lung Study (SLS) I[3] and II,[4] we hypothesized that those patients who experienced increased radiographic progression of ILD over the course of these trials would have worse long-term survival. The findings of this research may broaden our understanding of efficacy assessment in SSc-ILD research and ultimately shape the design of future clinical trials for this often-fatal disease.

STUDY DESIGN AND METHODS

Study participants

All participants enrolled in SLS I[3] (NCT01762449; NCT00004563) and SLS II[4] (NCT00883129) who had a follow up, standardized, high-resolution computed tomography (HRCT) scan of the chest were eligible to participate in this study. SLS I and II were randomized controlled trials (RCTs) that included an ethnically diverse population of both male and female patients with SSc-ILD followed at multiple centers across the United States. Eligibility criteria for these trials were similar.[3,4] The Institutional Review Board of each site approved the primary studies and long-term follow-up. Informed consent was obtained by all participants.

SLS I and II study design

In SLS I, 158 participants were randomized to receive oral cyclophosphamide (CYC) or placebo for 12 months and were followed for an additional 12 months off of therapy.[3] In SLS II,

142 patients were randomized to receive mycophenolate mofetil (MMF) for 24 months or oral CYC for 12 months followed by an additional 12 months of placebo.[4]

SLS I and II assessment measurements

The FVC (primary SLS I and II endpoint) was measured every 3 months during the 24-month study periods.[3,4] HRCT thoracic imaging was obtained at baseline and at the conclusion of active treatment in both trials (i.e., at 12 months in SLS I and at 24 months in SLS II). A Computer Aided Design scoring system[13-15] was used to calculate the Quantitative ILD (QILD) score for the whole lung at baseline and follow up (Please see e-Figure1 in Supplementary Material for further details on the HRCT imaging protocol and scoring). The QILD score included the sum of all abnormally classified scores, including fibrosis (e.g., reticular opacity with architectural distortion), ground glass opacity (e.g., increased parenchymal attenuation), and honeycombing (e.g., clustered air-filled cysts with dense walls).

The QILD threshold to define radiographic ILD progression (QILD score of ≥2%) was determined by first performing a statistical analysis of the cohorts using a quantitative imaging biomarker algorithm proposed in our prior publication.[17] This analysis examines both technical reproducibility and clinical reproducibility.[17] Technical reproducibility was reported as 0.60% in the limited agreement (Supplementary e-Figure 2). Subsequently, the factor of 2.77 ($=1.96*\sqrt{2}$) was multiplied by the variability of the two HRCT scans, resulting in a threshold of 1.66%. This number was then rounded up to 2%.

Next, we considered clinically meaningful correlates of a ≥2% increase in QILD, such as change in FVC%-predicted, patient reported-outcomes and mortality. In the SLS I and II cohorts,

an increase of QILD score of ≥2% was associated with a clinically meaningful decline in FVC% predicted.[18] Specifically, the majority of patients meeting the minimal clinically important difference (MCID) estimates for FVC%-predicted worsening also experienced an increase in QILD>2%. The MCID criteria for FVC worsening (-3.0% to -3.3%) were derived based on an analysis of two valid patient-reported outcomes.[18] As a final step, we tested the relationship between additional thresholds of QILD worsening (e.g., ≥1%, ≥3%, ≥4%, ≥5%) and survival, and the ≥2% increase in QILD demonstrated the strongest association with survival in both of the SLS I and II cohorts.

Long-term mortality

During the SLS I and II trials, mortality data were collected and causes of death were adjudicated by data safety and monitoring boards. After the 24-month trials, patients or their designated surrogates were contacted annually to assess morbidity and mortality outcomes. If the patient or his/her previously designated contact person could not be reached, investigators contacted site investigators and searched publicly available death registries (e.g., National Death Index and Social Security Death Index), as well as online obituaries. Survival status was ascertained for up to 12 and 8 years after the commencement of SLS I and II, respectively.

Statistical analysis

Baseline characteristics

Summary statistics were generated for baseline characteristics from the two cohorts. Group comparisons were performed using two-sample t-tests and chi-square tests.

Primary outcome: Survival

The primary outcome was all cause mortality. The Kaplan-Meier estimate was used to generate survival curves, and the log-rank test was used to compare survival between participants who experienced progression of ILD (≥2% increase in QILD score) versus those who experienced stability/improvement of ILD (<2% increase in QILD score). Cox proportional hazard models were subsequently developed. Using a comprehensive variable selection process described in our previous publication,[7] the following variables were found to be significantly associated with mortality in univariate analyses in both the SLS I and II cohorts: age, modified Rodnan skin score [mRSS], and baseline %-predicted values for FVC. The aforementioned variables were therefore included in the final Cox models. Treatment arm assignment was not significantly associated with mortality in SLS I or II;[7] however, given the potential interaction between treatment arm assignment and change in QILD, we created exploratory Cox models that also included treatment arm assignment. We also created exploratory Cox models that included variables associated with mortality based on expert knowledge (e.g., sex, race, diffuse SSc subtype, diffusing capacity for carbon monoxide [DLCO]%-predicted).

Secondary outcome: Course of FVC

Mixed effects models were created to compare the course of the FVC%-predicted (measured every 3 months over 24 months) between patients who experienced progression of ILD (≥2% increase in QILD score) versus those who experienced stability/improvement of ILD (<2% increase in QILD score). Covariates included the baseline FVC%-predicted, treatment arm assignment and the change in QILD score.

All tests were 2-sided and performed using SAS 9.4 (SAS Institute; Cary, NC).

RESULTS

Participant characteristics

Among all of the SLS I (N=158) and II (N=142) participants, 82 and 90 had follow up HRCT scans of the chest, respectively, and were included in the present analysis. In SLS I, the follow-up HRCT was part of a sub-study that was delayed in onset, limiting the ability to enroll patients who had already completed the 12-month follow-up visit. In SLS II, the primary reason not all patients had an HRCT scan of the chest at 24 months was due to participant drop-out and/or death.

The baseline disease features and demographic characteristics of the SLS I and II participants who had follow up HRCT scans were similar and closely reflected those of the overall study population (Supplementary e-Table 1). The majority of the patients were female with relatively early SSc (~3 years from the onset of the first non-Raynaud symptom attributable to SSc) and a moderate degree of restriction on pulmonary function testing.

Radiographic ILD progression in SLS I and II

In SLS I, 34 (41%) participants experienced an increase in QILD of ≥2% for the whole lung at 12 months. Among these participants, 13 were randomized to CYC and 21 were randomized to placebo. This difference between treatment arms is consistent with the statistically significant impact of CYC on the change in QILD score at 12 months reported in our prior studies.[12] In SLS II, 28 (31%) participants experienced an increase in QILD ≥2% for the whole lung at 24 months. Among these participants, 15 were randomized to CYC and 13 were randomized to MMF. In contrast to the outcome in SLS I, in which half of the participants

received placebo, no significant difference in treatment effect was observed between the two active treatment arms (MMF and CYC) in SLS II.[14] There were no significant differences in the baseline characteristics of participants who experienced an increase in QILD of ≥2% and those who did not in either SLS I or II, with the exception of a lower baseline DLCO%-predicted in SLS I participants who experienced an increase in QILD ≥2% (Table 1).

SLS I participants who experienced an increase in QILD ≥2% over 12 months were also more likely to experience a decline in the FVC%-predicted over the course of the trial; whereas SLS I participants who experienced a change in QILD <2% were more likely to experience stability/improvement in the FVC%-predicted (Figure 1a). Similarly, in SLS II, those participants who experienced an increase in QILD ≥2% over 24 months were also more likely to experience a decline in FVC%-predicted over the course of the trial; whereas those participants who experienced a change in QILD <2% were more likely to experience stability/improvement in FVC%-predicted (Figure 1b).

Long-term survival in SLS I and II

In SLS I, 66 (42%) of the entire cohort had died within 12 years after the first patient was randomized (CYC: 38; Placebo: 28). The majority (65%) of deaths were due to underlying SSc, and among the deaths due to SSc, 67% were due to respiratory failure. In the SLS I participants with follow up HRCT scans at 12 months, 28 (34%) deaths occurred during the long-term follow up period (CYC: 15; Placebo: 13).

In SLS II, 30 (21%) of the entire cohort had died within 8 years after the first patient was randomized (CYC: 16; MMF: 14). The majority (58%) of deaths were due to underlying SSc, and among the deaths due to SSc, 50% were due to respiratory failure. In the SLS II participants

with follow up HRCT scans at 24 months, 15 (17%) deaths occurred during the long-term follow up period (CYC: 6; MMF: 9).

Radiographic progression of ILD predicts survival

During the 12-year long-term follow up period, SLS I participants who experienced an increase in QILD of ≥2% for the whole lung at one year had a significantly increased risk of death (*P*= 0.01 by log-rank test; Figure 2). During the 8-year long-term follow up period, SLS II participants who experienced an increase in QILD of ≥2% for the whole lung at two years had a significantly increased risk of death (*P*= 0.019 by log-rank test; Figure 3).

After adjusting for age, mRSS and baseline %-predicted values for FVC, there was a suggestive association between increase in QILD of ≥2% and mortality in SLS I participants (HR 1.98, *P*=0.089; Table 2). In addition, after adjusting for the aforementioned variables, an increase in QILD of ≥2% was associated with a significantly increased risk of mortality in SLS II participants (HR 3.86, *P*=0.014; Table 3). After adding treatment arm assignment (Supplementary e-Tables 2 and 3), as well as the variables of sex, race, diffuse SSc subtype, DLCO%-predicted (Supplementary e-Tables 4 and 5) to the aforementioned Cox models, an increase in QILD of ≥2% remained associated with mortality in both SLS I (suggestive association) and SLS II (significant association).

In an exploratory analysis, physiologic progression of ILD was substituted for the baseline FVC%-predicted as a covariate in the Cox model. Physiologic progression was defined according to the OMERACT criteria as FVC%-predicted decline ≥10% *OR* FVC%-predicted decline between 5-9% *AND* DLCO%-predicted decline ≥15%.[6,19] In SLS I, 15 patients met the OMERACT criteria for physiologic ILD progression at 12 months. After adjusting for age,

mRSS and physiologic progression of ILD at 12 months, an increase in QILD of ≥2% remained significantly associated with an increased risk of mortality (HR 2.28, *P*=0.037; Supplementary e-Table 6). In another exploratory analysis, we included the change in FVC%-predicted at 12 months (measured continuously) as a covariate, and this covariate was not significantly associated with mortality (*P*=0.97; Supplementary e-Table 7).

In SLS II, too few patients (N= 3) patients met the OMERACT criteria for physiologic ILD progression at 24 months; therefore, a Cox model that included this covariate could not be created. However, similar to SLS I, we also created a Cox model that included the change in FVC%-predicted at 24 months (measured continuously) as a covariate, and there was a suggestive association between change in FVC%-predicted with mortality (HR 0.93, *P*=0.06; Supplementary e-Table 8).

DISCUSSION

Using data from two independent cohorts with extensive clinical characterization and follow up, the present study demonstrated that radiographic progression of SSc-ILD (defined by a ≥2 % increase in QILD score) over the course of one to two years is associated with an increased risk of long-term mortality. Even after adjusting for other factors known to affect survival in this patient population, worsening quantitative radiographic extent of ILD remained significantly associated with mortality in SLS II (suggestive association for SLS I).

Individualization of endpoints in RCTs for SSc-ILD is an evolving area of research with considerable implications for future drug discovery and development. Since no valid patient-reported outcomes currently exist for this disease state and an insufficient number of deaths

occur over standard trial periods,[3-5] there is an unmet need for establishing clinically meaningful and reliable surrogate endpoints for SSc-ILD RCTs.

The FVC has served as the primary endpoint for three of the largest SSc-ILD RCTs;[3-5] however, these trials all used different statistical methods for evaluating FVC change over time. For instance, in SLS I, a generalized estimating-equation regression model was created that used imputation for missing data to compare outcomes in the FVC%-predicted over one year.[3] In SLS II, a joint model, which combined a mixed effects model with a survival model to adjust for non-ignorable missing data due to drop-outs, treatment failures and deaths, was employed to compare the course of the FVC%-predicted over two years.[4] In the SENSCIS trial comparing nintedanib with placebo for SSc-ILD,[5] the annual rate of decline in FVC (milliliters per year) over one year was assessed with a random-coefficient regression model that also used imputation for missing data. In the recently published phase III trial of tocilizumab for the treatment of diffuse cutaneous SSc, FVC was a key secondary endpoint, and that study evaluated the difference in distribution of change from baseline to week 48 in the FVC%-predicted.[16] The diverse methodologies employed to evaluate how various treatments affect lung function render it difficult to compare treatment effects across studies. In addition, studies have demonstrated considerable fluctuations in declines in the FVC%-predicted within individual patients receiving treatment for ILD.[20] Finally, lung function outcomes correlated poorly with various patient reported outcomes in some trials,[5,21,22] raising the question as to whether lung function change is truly a clinically meaningful endpoint from the perspective of the patient.

The present study suggests that changes in the quantitative radiographic extent of ILD may serve as a proxy for long-term mortality in patients with SSc-ILD. Moreover, increased radiographic progression of ILD remained significantly associated with increased mortality in

the multivariable analysis in SLS II; whereas physiological progression on pulmonary function testing was not. These findings are consistent with studies of patients with other ILDs.[22]. For example, a recent study demonstrated that longitudinal changes in semi-quantitative visual HRCT fibrosis scores of >7% predicted lung transplantation-free survival in patients with idiopathic pulmonary fibrosis followed for an average of 3 years.[23] Moreover, progression of interstitial lung disease abnormalities in participants of the Framingham Heart Study based on visual assessment was associated with increased mortality.[24] However, the present study is the first study to demonstrate that increased *quantitative* radiographic progression of ILD beyond a pre-specified threshold over the course of both one- and two-years was associated with an increased risk of death in patients with SSc.

In addition to predicting survival, a prior study from SLS II demonstrated that changes in the quantitative radiographic extent of ILD in the whole lung were significantly associated with patient-reported outcomes, such as dyspnea.[22] Specifically, increased QILD over two years was associated with worse scores on the transitional dyspnea index.[22] Thus, radiographic progression of ILD may closely reflect actual changes in how a patient feels and functions.

In considering the results of the multivariable analyses, the hazard ratio for change in QILD was higher for the SLS II survival model compared with the SLS I model. This finding could suggest that radiographic progression measured at two years is a better predictor of long-term mortality than one year, although to confirm this hypothesis, one would need to measure QILD changes at these two time points in the same cohort. It is also notable that even though substantially more patients randomized to placebo experienced radiographic progression compared with CYC, the long-term mortality rates were similar between the SLS I study

treatment arms, suggesting that one year of therapy is unlikely to lead to a sustained improvement in long-term survival.

Interestingly, the baseline characteristics of study participants who experienced radiographic progression of ILD and those who did not were fairly similar. Although not statistically significant, a greater proportion of males and African American participants experienced radiographic progression of ILD during these studies. These analyses may have been underpowered to detect significant differences, and larger studies are needed to determine whether sex and race affect radiographic progression of SSc-ILD.

While quantitative imaging analysis can more sensitively detect changes in the radiographic extent of ILD than visual assessment,[25] a major limitation of this study is that automated methods for radiographic ILD scoring are not yet widely available in clinical practice. However, this approach has been used in a number of SSc RCTs,[3,4,16] demonstrating its feasibility as a study endpoint. Furthermore, insufficient inspiration and/or respiratory muscle strain may curtail lung expansion during HRCT assessment. Furthermore, superimposed infection may also confound quantitative image assessment. Reassuringly, study technicians were trained to ensure adequate inspiration, and all images underwent quality control assessment to evaluate for the presence of infection prior to quantitative image analysis. Another potential short-coming of this study was that radiographic assessment was performed at different time points in SLS I (one year) and II (two years), which could introduce lead time bias. However, few deaths occurred during the active treatment periods (N=5 and 16 for SLS I and II, respectively). Moreover, this could also be perceived as a strength as the findings suggest that radiographic changes at either of these time points have prognostic value.

Additional strengths of this study include the evaluation of two multi-center SSc-ILD cohorts who received standard treatment and follow up during the radiographic assessment period. Moreover, while not all participants had a follow up HRCT assessment, more than half of all patients did, and there was no difference in the baseline characteristics of participants without follow up HRCT assessment and the entire study cohort. Finally, to our knowledge, this is the longest time period participants in any RCT for SSc-ILD have been followed for mortality outcomes.

INTERPRETATION

In summary, measuring changes in the quantitative radiographic extent of ILD predicts long-term mortality in patients with SSc. Moreover, radiographic progression of ILD appeared to be a stronger predictor of mortality than longitudinal functional decline in two independent, multi-center cohorts. Radiographic endpoints may serve as more reliable and reproducible endpoints in SSc-ILD trials compared with the FVC, and this is the first study to suggest that these endpoints are surrogates for mortality. Future studies are needed to determine whether measuring the change in QILD at earlier time points (e.g., 6 months) could be used to detect treatment effects and predict long-term outcomes in these patients.

TAKE-HOME POINT

Study Question: Do short-term changes in radiographic measures of ILD predict long-term survival in patients with SSc?

Results: In SSc-ILD patients enrolled in two large, randomized controlled trials, an increase in the quantitative radiographic extent of ILD of ≥2% over 1 to 2 years was associated with worse long-term survival.

Interpretation: Short-term changes in the radiographic extent of ILD in SSc patients receiving treatment or placebo may serve as a proxy for long-term mortality. Future SSc-ILD clinical trials should consider including radiographic endpoints to assess treatment response.

ACKNOWLEDGEMENTS

Guarantor statement: ERV takes full responsibility for the content of the manuscript, including the data and analysis.

Author contributions:

ERV: Substantial contributions to the conception and design of the work, the acquisition, analysis, or interpretation of data for the work; Drafting the work and revising it critically for important intellectual content; Final approval of the version to be published; Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

DPT: Substantial contributions to the conception and design of the work, the acquisition, analysis, or interpretation of data for the work; Revising the work critically for important intellectual content; Final approval of the version to be published; Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

MDR: Substantial contributions to the conception and design of the work, the acquisition, analysis, or interpretation of data for the work; Revising the work critically for important intellectual content; Final approval of the version to be published; Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

JG: Substantial contributions to the conception and design of the work, the acquisition, analysis, or interpretation of data for the work; Revising the work critically for important intellectual content; Final approval of the version to be published; Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

GHJK: Substantial contributions to the conception and design of the work, the acquisition, analysis, or interpretation of data for the work; Revising the work critically for important intellectual content; Final approval of the version to be published; Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Financial/nonfinancial disclosures:

ERV: Consulting and speaker fees from Boehringer Ingelheim; Grant support from Kadmon, Horizon Therapeutics

DPT: Grant support from Genentech/Roche

MDR: Grant support from Genentech/Roche

JG: None

GHJK: None

Role of the funding sources: The study sponsors provided funding for study materials and data analysis. They had no role in the study design; collection, analysis, and the interpretation of the data; in the writing of the report; and in the decision to submit the paper for publication.

Other contributions:

The following persons and institutions participated in the Scleroderma Lung Study I: *University of California at Los Angeles (UCLA), Los Angeles*: P.J. Clements, D.P. Tashkin, R. Elashoff, J. Goldin, M. Roth, D. Furst, K. Bulpitt, D. Khanna, W.-L.J. Chung, S. Viasco, M. Sterz, L. Woolcock, X. Yan, J. Ho, S. Vasunilashorn, I. da Costa; *University of Medicine and Dentistry of New Jersey, New Brunswick*: J.R. Seibold, D.J. Riley, J.K. Amorosa, V.M. Hsu, D.A. McCloskey, J.E. Wilson; *University of Illinois Chicago, Chicago*: J. Varga, D. Schraufnagel, A. Wilbur, M. Lopata, S. Arami, P. Cole-Saffold; *Boston University, Boston*: R. Simms, A. Theodore, P. Clarke, J. Korn, K. Tobin, M. Nuite; *Medical University of South Carolina, Charleston*: R. Silver, M. Bolster, C. Strange, S. Schabel, E. Smith, J. Arnold, K. Caldwell, M. Bonner; *Johns Hopkins School of Medicine, Baltimore*: R. Wise, F. Wigley, B. White, L. Hummers, M. Bohlman, A. Polito, G. Leatherman, E. Forbes, M. Daniel; *Georgetown University, Washington, D.C.*: V. Steen, C. Read, C. Cooper, S. Wheaton, A. Carey, A. Ortiz; *University of Texas at Houston, Houston*: M. Mayes, E. Parsley, S. Oldham, T. Filemon, S. Jordan, M. Perry; *University of California at San Francisco, San Francisco*: K. Connolly, J. Golden, P. Wolters, R. Webb, J. Davis, C. Antolos, C. Maynetto; *University of Alabama at Birmingham, Birmingham*: B. Fessler, M. Olman, C. Sanders, L. Heck, T. Parkhill; *University of Connecticut Health Center, Farmington*: N. Rothfield, M. Metersky, R. Cobb, M. Aberles, F. Ingenito, E. Breen; *Wayne State University, Detroit*: M. Mayes, K. Mubarak, J.L. Granda, J.

Silva, Z. Injic, R. Alexander; *Virginia Mason Research Center, Seattle*: D. Furst, S. Springmeyer, S. Kirkland, J. Molitor, R. Hinke, A. Mondt; *Data Safety and Monitoring Board*: Harvard Medical School, Boston — T. Thompson; Veterans Affairs Medical Center, Brown University, Providence, R.I. — S. Rounds; Cedars Sinai–UCLA, Los Angeles — M. Weinstein; Clinical Trials Surveys, Baltimore — B. Thompson; *Mortality and Morbidity Review Committee*: UCLA, Los Angeles — H. Paulus, S. Levy; Johns Hopkins University, Baltimore — D. Martin.

The following persons and institutions participated in the Scleroderma Lung Study II: *University of Boston, Boston:* A.C. Theodore, R.W. Simms, E. Kissin, F.Y. Cheong; *Georgetown University, Washington, D.C.*: V.D. Steen, C.A. Read Jr., C. Fridley, M. Zulmatashvili; *Johns Hopkins University, Baltimore*: R.A. Wise, F.M. Wigley, L. Hummers, G. Leatherman; *Medical University of South Carolina, Charleston*: R.M. Silver, C. Strange, F.N. Hant, J. Ham, K. Gibson, D. Rosson; *University of California, Los Angeles (UCLA), Los Angeles*: D.P. Tashkin, R.M. Elashoff, M.D. Roth, P.J. Clements, D. Furst, S. Kafaja, E. Kleerup, D. Elashoff, J. Goldin, E. Ariola, G. Marlis, J. Mason-Berry, P. Saffold, M. Rodriguez, L. Guzman, J. Brook; *University of California, San Francisco (UCSF), San Francisco*: J. Golden, M.K. Connolly, A. Eller, D. Leong, M. Lalosh, J. Obata; *University of Illinois, Chicago*: S. Volkov, D. Schraufnagel, S. Arami, D. Franklin; *Northwestern University, Chicago*: J. Varga, J. Dematte, M. Hinchcliff, C. DeLuca, H. Donnelly, C. Marlin; *University of Medicine and Dentistry of New Jersey, New Brunswick*: D.J. Riley, V.M. Hsu, D.A. McCloskey; *University of Michigan, Ann Arbor*: K. Phillips, D. Khanna, F.J. Martinez, E. Schiopu, J. Konkle; *University of Texas, Houston*: M. Mayes, B. Patel, S. Assassi, F. Tan; *National Jewish Health, Denver*: A. Fischer, J. Swigris, R. Meehan, K. Brown, T. Warren, M. Morrison; *University of Utah, Salt Lake City*: M. B.

Scholand, T. Frecht, P. Carey, M. Villegas; *University of Minnesota*, *Minneapolis*: J. Molitor, P. Carlson.

REFERENCES

1. 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.

2. Elhai M, Meune C, Boubaya M, et al., EUSTAR group. Mapping and predicting mortality from systemic sclerosis. *Ann Rheum Dis.* 2017;76:1897-1905.

3. Tashkin DP, Elashoff R, Clements PJ, et al., for the Scleroderma Lung Study Research Group. Cyclophosphamide versus placebo in scleroderma lung disease. *N Engl J Med*. 2006;354:2655-2666.

4. Tashkin DP, Roth MD, Clements PJ, et al., for the Scleroderma Lung Study II Investigators. Mycophenolatge mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease: Scleroderma Lung Study II (SLS II), a double-blind, parallel group, randomized controlled trial. *Lancet Respir Med*. 2016;4:708-719.

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

6. Goh NS, Hoyles RK, Denton CP, et al. Short-term pulmonary function trends are predictive of mortality in interstitial lung disease associated with systemic sclerosis. *Arthritis Rheum*. 2017;69:1670-1678.

7. Volkmann ER, Tashkin DP, Sim M, et al., SLS I and SLS II study groups. Short-term progression of interstitial lung disease in systemic sclerosis predicts long-term survival in two independent clinical trial cohorts. *Ann Rheum Dis.* 2019;78:122-130.

8. Nawata T, Shirai Y, Suzuki M, Kuwana M. Chest wall muscle atrophy as a contributory factor for forced vital capacity decline in systemic sclerosis-interstitial lung disease. *Rheumatology (Oxford).* 2021;60:250-255.

9. Tashkin DP, Volkmann ER, Tseng C-H, et al. Relationship between quantitative radiographic assessments of interstitial lung disease and physiologic and clinical features of systemic sclerosis. *Ann Rheum Dis.* 2016;75:374-381.

10. Camiciottoli G, Orlandi I, Bartolucci M, et al. Lung CT densitometry in systemic sclerosis: correlation with lung function, exercise testing, and quality of life. *Chest*. 2007;131:672-681.

11. Khanna D, Nagaraja V, Tseng CH, et al. Predictors of lung function decline in scleroderma-related interstitial lung disease based on high-resolution computed tomography: implications for cohort enrichment in systemic sclerosis-associated interstitial lung disease trials. *Arthritis Res Ther.* 2015;17:372.

12. Roth MD, Tseng CH, Clements PJ, et al. Predicting treatment outcomes and responder subsets in scleroderma-related interstitial lung disease. *Arthritis Rheum*. 2011;63:2797-808.

13. Kim HJ, Brown M, Elashoff F, et al. Quantitative texture-based assessment of one-year changes in fibrotic reticular patterns on HRCT in scleroderma lung disease treated with oral cyclophosphamide. *Eur Radiology.* 2011;21:2455-2465.

14. Kim HJ, Tashkin DP, Chong S, et al. Transitions to different patterns of interstitial lung disease in scleroderma with and without treatment. *Ann Rheum Dis*. 2016; 75:1367-1371.

15. Goldin JG, Kim GH, Volkmann ER, et al. Longitudinal changes in quantitative lung disease on CT after immunosuppression in the Scleroderma Lung Study II. *Ann Am Thorac Soc*. 2018;15:1286-1295.

16. Khanna D, Lin CJF, Furst DE, et al., focuSSced investigators. Tocilizumab in systemic sclerosis: a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Respir Med*. 2020;8:963-974.

17. Obuchowski NA, Reeves AP, Huang EP, et al; Algorithm Comparison Working Group. Quantitative imaging biomarkers: a review of statistical methods for computer algorithm comparisons. *Stat Methods Med Res*. 2015;24:68-106.

18. Kafaja S, Clements P. Wilhalme H, et al. Reliability and minimally clinically important differences of forced vital capacity: Results from the Scleroderma Lung Studies (SLS-I and SLS-II). *Am J Respir Crit Care Med*. 2018;197:644-652.

19. Khanna D, Mittoo S, Aggarwal R, et al. Connective tissue disease-associated interstitial lung diseases (CTD-ILD) - Report from OMERACT CTD-ILD Working Group. *Journal of Rheumatology*. 2015; **42**: 2168-2171.

20. Nathan SD, Albera C, Bradford WZ, et al. Effect of continued treatment with pirfenidone following clinically meaningful declines in forced vital capacity, analysis of data from three phase 3 trials in patients with idiopathic pulmonary fibrosis. *Thorax.* 2016;71:429–435.

21. Khanna D, Yan X, Tashkin DP, et al., for the Scleroderma Lung Study Group. Impact of oral cyclophosphamide on health-related quality of life in patients with active scleroderma lung disease: Results from the Scleroderma Lung Study. *Arthritis Rheum*. 2007;56:1676-1684.

22. Volkmann ER, Tashkin DP, LeClair H, et al. Treatment with mycophenolate mofetil and cyclophosphamide leads to clinically meaningful improvements in patient-reported outcomes in scleroderma lung disease: Results of Scleroderma Lung Study II. *ACR Open Rheum*. 2020;2:362-370.

23. Taha N, D'Amato D, Hosein K, et al. Longitudinal functional changes with clinically significant radiographic progression in idiopathic pulmonary fibrosis: are we following the right parameters. *Respiratory Research*. 2020;21:119.

24. Araki T, Putman RK, Hatabu H, et al. Development and progression of interstitial lung abnormalities in the Framingham Heart Study. *Am J Respir Crit Care Med*. 2016;194:1514-1522

25. 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:S26-35.

TABLES

Table 1. Baseline patient characteristics of SLS I and II participants who had ΔQILD ≥2%versus ΔQILD<2% at 12 and 24 months, respectively

	SLS I ΔQILD ≥2% (N=34)	SLS I ΔQILD<2% (N=48)	P-value	SLS II ΔQILD ≥2% (N=25)	SLS II ΔQILD<2% (N=65)	P-value
Age (years)- Mean (SD)	48.1 (11.2)	45.6 (11.7)	0.27	51.63 (9.74)	51.36 (8.98)	0.81
Female- %	21 (61.76%)	39(81.25%)	0.050	15 (60.00%)	51 (78.46%)	0.053
SSc Duration (years)- Median (IQR)[a]	2.18 (2.42)	3.10 (3.47)	0.031	2.50 (2.83)	1.67 (3.00)	0.59
Diffuse- N (%)	21 (61.76%)	27 (56.25%)	0.62	15 (60.00%)	38 (58.46%)	0.96
Race- N (%)[b]						
White	21 (61.76%)	34 (70.83%)	0.62	15 (60.00%)	44 (67.69%)	0.32
African American	6 (17.65%)	5 (10.42%)		8 (32.00%)	14 (21.54%)	
Asian	2 (5.88%)	1 (2.08%)		1 (4.00%)	5 (7.69%)	
Other	5 (14.71%)	7 (14.58%)		1 (4.00%)	2 (3.08%)	
Unknown	0 (0.00%)	1 (2.08%)		0 (0.00%)	0 (0.00%)	
MRSS- Mean (SD)	16.53 (11.63)	14.15 (10.78)	0.33	12.93 (7.92)	14.20 (10.05)	0.82
History of prior smoking N (%)	11 (33.33%)	17 (35.42%)	0.85	6 (25.00%)	19 (29.23%)	0.69

FVC % Predicted- Mean (SD)	66.67 (10.85)	71.16 (11.68)	0.21	67.69 (7.78)	65.91 (9.00)	0.93
DLCO % Predicted- Mean (SD)[b]	43.20 (12.70)	50.46 (14.54)	0.022	55.94 (13.91)	55.42 (12.74)	0.87
QLF % Whole Lung- Mean (SD)	8.23 (5.98)	10.14 (10.78)	0.35	7.77 (5.72)	7.80 (7.01)	0.97
QILD % Whole Lung- Mean (SD)	31.57 (11.99)	35.62 (17.04)	0.24	26.61 (8.84)	27.05 (13.91)	0.86

[a] N=81 and N=84 for disease duration in SLS I and II, respectively.

[b] N=81 for race in SLS I.

Table 2. Cox Proportional Hazards model for survival in SLS I (N=82)[a]

	Hazard ratio	95% CI	P-value	C-index
Univariate Analyses				
ΔQILD ≥2%	2.61	(1.23, 5.57)	0.013	0.637
Continuous change in QILD	1.02	(0.99, 1.06)	0.16	0.643
Multivariable Analyses				
Age	1.04	(1.01, 1.08)	0.019	0.728
mRSS	1.04	(1.00, 1.07)	0.028	
Baseline FVC%-predicted	0.97	(0.93, 1.01)	0.098	
ΔQILD ≥2%	1.98	(0.90, 4.40)	0.089	

Table 3. Cox Proportional Hazards model for survival in SLS II (N=90)[a]

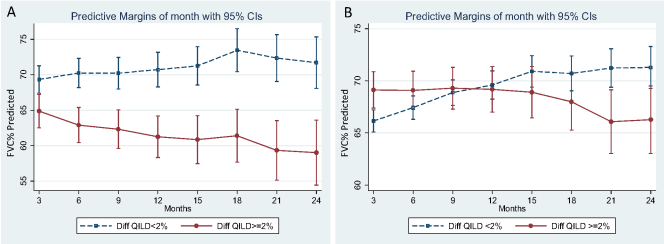
	Hazard ratio	95% CI	P-value	C-index
Univariate Analyses				
ΔQILD ≥2%	3.17	(1.15, 8.76)	0.026	0.688
Continuous change in QILD	1.10	(1.03, 1.17)	0.004	0.735
Multivariable Analyses				
Age	1.05	(1.00, 1.11)	0.074	0.743
mRSS	1.00	(0.94, 1.07)	0.614	
Baseline FVC%-predicted	1.01	(0.95, 1.06)	0.895	
ΔQILD ≥2%	3.86	(1.31, 11.27)	0.014	

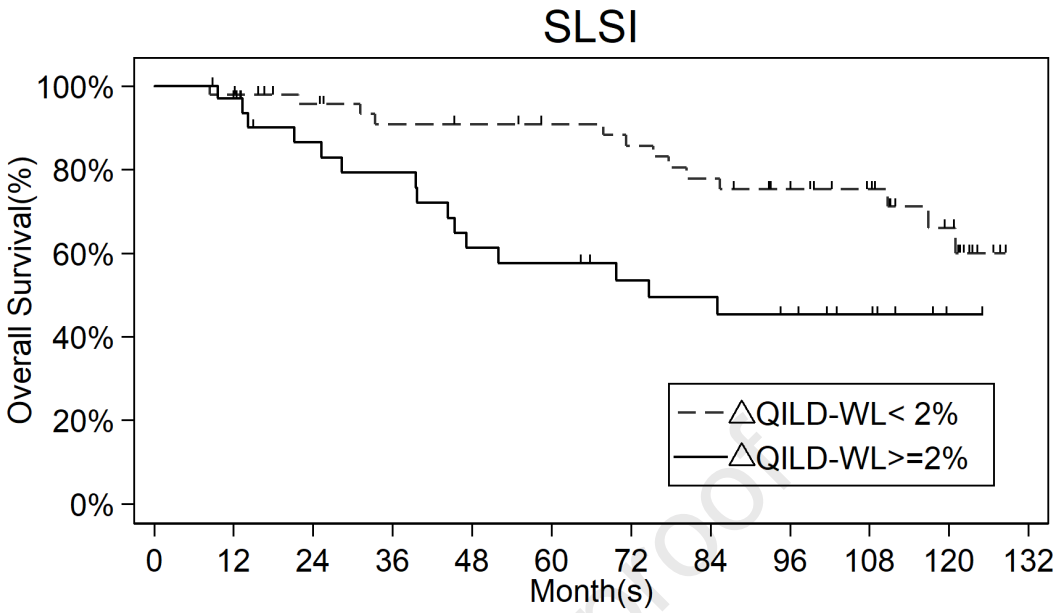
FIGURE LEGENDS

Figure 1. Mixed effects model of the course of the FVC%-predicted over 24 months for participants with ≥2% (red line) versus <2% (blue line) increase in QILD score for SLS I (Figure 1a) and SLS II (Figure 1b). Covariates for both models included the baseline FVC% predicted and treatment arm.

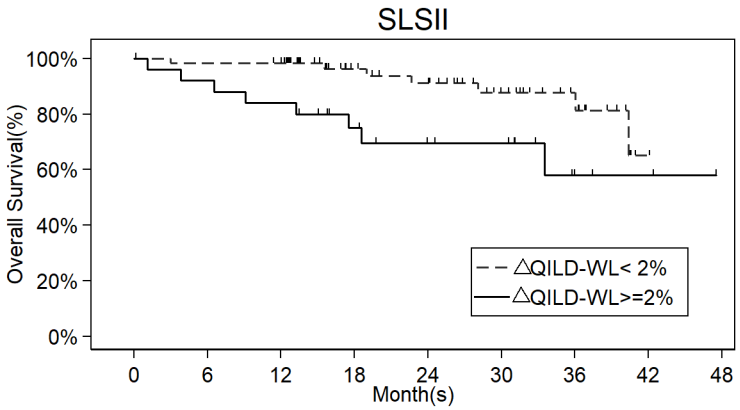
Figure 2. Kaplan-Meier plot for overall survival for participants with ≥2% versus <2% increase in QILD score for the whole lung from baseline to 12 months in SLS I.

Figure 3. Kaplan-Meier plot for overall survival for participants with ≥2% versus <2% increase in QILD score for the whole lung from baseline to 24 months in SLS II.





Number at risk												
△QILD-WL< 2%	48	46	42	38	37	35	33	30	24	20	12	0
△QILD-WL>=2%	34	31	24	22	17	16	13	12	10	6	1	0



Number at risk									
△QILD-WL< 2%	65	63	61	39	33	22	14	0	0
△QILD-WL>=2%	25	23	21	15	11	10	3	2	0