CLINICAL SCIENCE

Short-term progression of interstitial lung disease in systemic sclerosis predicts long-term survival in two independent clinical trial cohorts

Elizabeth R Volkmann,[1] Donald P Tashkin,[1] Myung Sim,[1] Ning Li,[2] Ellen Goldmuntz,[3] Lynette Keyes-Elstein,[4] Ashley Pinckney,[4] Daniel E Furst,[1] Philip J Clements,[1] Dinesh Khanna,[5] Virginia Steen,[6] Dean E Schraufnagel,[7] Shiva Arami,[7] Vivien Hsu,[8] Michael D Roth,[1] Robert M Elashoff,[2] Keith M Sullivan,[9] SLS I and SLS II study groups

ABSTRACT

Objective To assess survival and identify predictors of survival in patients with systemic sclerosis-interstitial lung disease (SSc-ILD) who participated in the Scleroderma Lung Studies (SLS) I and II.

Methods SLS I randomised 158 patients with SSc-ILD to 1 year of oral cyclophosphamide (CYC) vs placebo. SLS II randomised 142 patients to 1 year of oral CYC followed by 1 year of placebo vs 2 years of mycophenolate mofetil. Counting process Cox proportional hazard modelling identified variables associated with long-term mortality in SLS I and II. Internal validation was performed using joint modelling.

Results After a median follow-up of 8 years, 42% of SLS I patients died, and when known the cause of death was most often attributable to SSc. There was no significant difference in the time to death between treatment arms in SLS I or II. Higher baseline skin score, older age, and a decline in the forced vital capacity (FVC) and the diffusing capacity for carbon monoxide (DLCO) over 2 years were independently associated with an increased risk of mortality in SLS I. The Cox model identified the same mortality predictor variables using the SLS II data.

Conclusion In addition to identifying traditional mortality risk factors in SSc (skin score, age), this study demonstrated that a decline in FVC and DLCO over 2 years was a better predictor of mortality than baseline FVC and DLCO. These findings suggest that short-term changes in surrogate measures of SSc-ILD progression may have important effects on long-term outcomes.

Key messages

What is already known about this subject?
► Observational studies have identified factors associated with an increased risk of mortality in systemic sclerosis (SSc), including age, extent of cutaneous sclerosis and severity of interstitial lung disease (ILD).

What does this study add?
► In addition to identifying traditional risk factors for mortality, this study found that short-term progression of ILD was a better predictor of mortality than baseline severity of ILD.
► Specifically, patients who experienced a decline in lung function and diffusing capacity over 2 years had a substantially increased risk of mortality even after adjusting for treatment arm assignment, as well as baseline disease severity.

How might this impact on clinical practice or future developments?
► Patients with SSc-ILD who experience an early decline in lung function and have an increased risk of death may benefit from receiving a more aggressive treatment approach that could include escalation of immunosuppression, addition of antifibrotic therapy and/or evaluation for haematopoetic stem cell transplant.
► SSc providers should closely monitor lung function when ILD is present to accurately identify declines in lung function and promptly intervene to improve patient outcomes.

INTRODUCTION

Interstitial lung disease (ILD) is the leading cause of death in systemic sclerosis (SSc),[1] accounting for over one-third of SSc-related deaths in a multicentre observational study of over 5000 patients with SSc.[2] In addition, ILD occurs in the majority of patients with SSc[3] and is found in 79% of patients with SSc at autopsy.[4]

Immunosuppressant agents, such as mycophenolate motetil (MMF) and cyclophosphamide (CYC), are currently used to treat SSc-ILD.[5–7] However, randomised controlled trials (RCTs) have demonstrated that while some patients experience improvement in lung function after treatment with MMF or CYC, other patients experience ILD progression despite treatment with immunosuppression.[6 7] Moreover, not all patients with ILD develop symptoms or will have progressive disease that leads to death even in the absence of treatment.[3 8]

The present study sought to develop a mortality prediction model using data from two RCTs for SSc-ILD (Scleroderma Lung Study (SLS) I and II).[6 7] Using data from RCTs (in contrast to an observational cohort) may minimise confounding due to factors that affect survival, such as timing of treatment initiation, access to healthcare, socioeconomic

status as well as comorbid conditions. From these two RCTs with rigorous entry criteria, close monitoring of pulmonary function every 3 months for 2 years and standard treatment regimens, the hypothesis was that in this controlled setting of 300 participants with SSc and ILD new predictors of mortality could be discovered.

METHODS

Study participants

All participants enrolled in SLS I[6] (NCT01762449; NCT00004563) and SLS II[7] (NCT00883129) were eligible to participate in the SLS long-term follow-up study. SLS I and II included adult patients with SSc with evidence of ILD on high-resolution CT (HRCT) with a duration of disease ≤7 years from onset of the first non-Raynaud's symptom of SSc. (Please see online supplementary appendix for complete eligibility criteria.) Only participants who provided informed consent were included in the present analyses.

SLS I and II study design

In SLS I, 158 participants were randomised to receive either oral CYC or matching placebo for 1 year, followed by an additional year of observation off-treatment as previously published.[6] In SLS II, 142 patients were randomised to receive either MMF for 2 years or oral CYC for 1 year, followed by an additional year placebo using a double-dummy design to maintain the blinding as reported.[7]

SLS I and II assessment measurements

Complete details of SLS I and II assessment measurements are in the online supplementary appendix. Forced vital capacity (FVC) (primary SLS I and II endpoint) and diffusing capacity for carbon monoxide (DLCO) (secondary SLS I and II endpoint) were measured every 3 months during the 24-month study period for both studies.[6 7] HRCT thoracic imaging was obtained at baseline and 24 months in SLS II and at baseline and 12 months in SLS I. Quantitative imaging analysis (to quantify the extent of ILD) was performed as previously published and is described in the online supplementary appendix.[32]

Long-term follow up assessment

During both the SLS I and II study periods, when the statistical centre was informed of a participant's death, clinical research associates were asked to collect source documentation to determine the cause of death. A mortality and morbidity committee adjudicated the causes of death to determine whether the cause was related to underlying SSc, medication or another cause based on expert consensus. Following the 24-month study periods, patients or designated surrogates were contacted to assess morbidity and mortality outcomes. (Please see online supplementary appendix for further details of the long-term follow-up assessment.)

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 χ^2 tests.

Primary outcome: survival

The primary outcome was survival. The Kaplan-Meier estimate was used to generate survival curves, and the log-rank test was used to compare survival between groups. If survival status was unknown, survival time was censored at the date when the participant was last known to be alive. Cox proportional hazard models were developed to evaluate the impact of covariates shown previously to be associated with survival, including treatment, baseline Modified Rodnan Skin Score (MRSS), age, sex, race, disease duration, type of SSc (limited or diffuse), serological subtype (Scl-70 antibody-positive, RNA polymerase III antibody-positive), % predicted values for FVC or DLCO, and the radiographic quantitative extent of ILD/fibrosis. Models for FVC and DLCO were first fit using the baseline measure as the covariate of interest, and then in separate models using the longitudinal assessments over 24 months as a time-varying covariate. Final models were validated by fitting joint models for longitudinal and survival data using the SAS macro, JMFit[9] (see online supplementary appendix for details on variables' definitions and selection and joint modelling methods).

All tests were two-sided. All analyses were performed using SAS V.9.4.

RESULTS

Participant characteristics

The baseline characteristics of the participants in SLS I and SLS II were fairly similar (table 1). SLS II participants were slightly older and had shorter disease durations compared with SLS I participants. While the FVC%-predicted did not differ between the two cohorts, SLS II participants had slightly more restrictive ventilatory impairment, as reflected by lower total lung capacity, despite less diffusion impairment.

Participant disposition

SLS I

Twelve years after the first patient was randomised in SLS I, 66 of 158 (42%) participants had died (CYC: 38; placebo 28). Among the 37 patients for whom the cause of death was known, 24 deaths (65%) were attributable to underlying SSc, of which 16 (CYC 8; placebo: 8) were due to respiratory failure (table 2). Two of the deaths (1 CYC, 1 placebo) due to 'Respiratory Failure' were not attributed to underlying SSc. Survival status could not be determined in 34 participants. The median follow-up time for all patients in SLS I was 8 years.

SLS II

Eight years after the first patient was randomised in SLS II, 30 of 142 (21%) participants had died (CYC: 16; MMF: 14). Among the 26 patients for whom the cause of death was known, 15 deaths (58%) were attributable to underlying SSc, of which 13 (CYC: 6; MMF: 7) resulted from respiratory failure (table 2). Survival status could not be determined in 12 participants. The median follow-up time for all patients in SLS II was 3.6 years.

CYC does not improve long-term survival compared with placebo in SLS I

During the 24-month study period of SLS I, six participants randomised to CYC and six participants randomised to placebo expired.[10] During the 12-year long-term follow-up period, there was no significant difference in the time to death (p=0.335 by log-rank test; figure 1A) nor the time to death or organ failure (p=0.539 by log-rank test; figure 1B) for patients randomised to CYC versus placebo in SLS I. Moreover, time to the development of organ failure did not differ between the two study arms (p=0.185 by log-rank test; online supplementary figure S1), nor did time to the development of malignancy (p=0.701 by log-rank test; online supplementary figure S2). Types/locations of malignancies in SLS

Systemic sclerosis

Table 1 Baseline characteristics of SLS I and SLS II participants

Measure	SLS I		SLS II		P values*
	n	Mean±SD (CYC; placebo)	n	Mean±SD (CYC; MMF)	(SLS I vs SLS II)
Age, years	158	48.5±12.3 (48.4±12.3; 47.7±12.5) P=0.693	142	52.3±9.7 (52.0±9.8; 52.6±9.7) P=0.725	0.004
Gender (female/male) (%)	158	70/30 (76/24; 65/35) P=0.117	142	74/26 (78/22; 70/30) P=0.248	0.477
Race (%)					
White	158	67 (67; 66)	142	68 (66; 71)	
African–American		16 (15; 18)		23 (26; 20)	
Asian		4 (4; 5)		6 (4; 9)	
Other		12 (14; 10)		1 (3; 0)	
Unknown		1 (0; 1) P=0.776		0 P=0.324	
Diffuse %/Limited %	158	59/41 (62/38; 57/43) P=0.512	142	59/41 (55/45; 62/38) P=0.363	0.855
Disease duration, years	158	3.2±2.1 (3.2±2.3; 3.1±1.8) P=0.867	139	2.6±1.8 (2.5±1.8; 2.6±1.7) P=0.757	0.008
FVC, % predicted	158	68.1±12.1 (67.6±11.3; 68.6±13.0) P=0.640	142	66.5±9.1 (66.5±9.9; 66.5±8.3) P=1.000	0.194
FEV_1/FVC, %	155	82.8±8.0 (82.8±5.8; 82.8±7.4) P=0.999	142	82.6±5.6 (83.3±5.6; 81.8±5.5) P=0.105	0.802
TLC, % reference	156	69.6±13.1 (69.8±12.9; 69.3±13.3) P=0.845	142	65.9±10.9 (65.5±12.0; 66.3±10.0) P=0.668	0.008
DLCO, % reference	158	47.2±14.0 (47.1±13.7; 47.4±14.3) P=0.881	142	54.0±12.7 (54.1±14.1; 54.0±11.1) P=0.963	<0.001
BDI (focal score, 0–12)†	156	5.7±1.8 (5.6±1.7; 5.7±2.0) P=0.772	134	7.1±2.2 (7.1±2.3; 7.3±2.1) P=0.537	<0.001
HAQ-DI (score, 1–3)‡	157	0.83±0.7 (1.0±0.7; 0.7±0.7) P=0.018	142	0.72±0.7 (0.7±0.7; 0.7±0.6) P=0.784	0.176
VAS – breathlessness (0–100)	156	28.4±26.1 (27.2±24.6; 29.5±27.6) P=0.581	139	24.5±28.1 (24.4±28.9; 24.5±27.4) P=0.983	0.220
MRSS (0–51)	158	14.8±10.9 (15.6±11.4; 14.0±10.5) P=0.361	142	14.7±10.5 (14.0±10.6; 15.3±10.4) P=0.471	0.936
Lung fibrosis (QLFib) score, whole lung (WL), %	125	10.2±10.4 (10.3±10.5; 10.1±10.4) P=0.904	137	8.6±6.9 (8.9±7.0; 8.3±6.8) P=0.578	0.148
Lung fibrosis (QLFib) score, worst zone (ZM), %	125	26.5±21.9 (28.2±23.4; 25.0±20.5) P=0.407	137	22.8±19.6 (22.6±19.3; 23.0±20.2) P=0.903	0.152
Quantitative ILD (QILD) score, % WL	125	35.5±16.9 (35.8±17.1; 35.3±16.9) P=0.855	137	29.5±14.0 (31.6±14.4; 27.2±13.2) P=0.064	0.002
Quantitative ILD (QILD) score, % ZM	125	58.1±21.7 (58.1±22.3; 58.0±21.3) P=0.977	137	51.2±20.3 (52.3±19.9; 50.0±20.9) P=0.517	0.009

BDI, Baseline Dyspnoea Index; CYC, cyclophosphamide; DLCO, single-breath diffusing capacity for carbon monoxide; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; HAQ-DI, Health Assessment Questionnaire for Scleroderma-Disability Index; HRCT, high-resolution CT; MMF, mycophenolate mofetil; MRSS, Modified Rodnan Skin Score; QLFib WL, %, quantitative extent of lung fibrosis (reticulations) in whole lung on high-resolution CT; QLFib ZM, %, quantitative extent of lung fibrosis in the zone of maximal involvement on HRCT; QILD WL, %, quantitative extent of interstitial lung disease (fibrosis+ground class opacity (GGO)+honeycombing) in whole lung on HRCT; QILD ZM, quantitative extent of interstitial lung disease (fibrosis+GGO+honeycombing) in the zone of maximal involvement on HRCT; SLS, Scleroderma Lung Studies; TLC, total lung capacity; VAS, Visual Analogue Scale.

*t-Tests were used for all comparisons with the exception of gender and diffuse vs limited disease (χ^2 test).

†High score denotes worse dyspnoea.

‡High score denotes worse function.

Table 2 Long-term morbidity and mortality outcomes of SLS I and II participants

	SLS I		SLS II	
	CYC (n=79)	Placebo (n=79)	CYC (n=73)	MMF (n=69)
Subject status				
Alive	29 (37%)	29 (37%)	50 (68%)	51 (74%)
Lost to follow-up*	10 (13%)	22 (28%)	7 (10%)	4 (6%)
No data available	2 (3%)	0	0	0
Deceased	38 (48%)	28 (35%)	16 (22%)	14 (20%)
Death related to SSc	12	12	7	8
Death unrelated to SSc	6	4	8	3
Unknown if related to SSc	20	12	1	3
Cause of death†				
Respiratory failure (ILD)‡	9	9	6	7
Aspiration	0	0	2	0
Pulmonary hypertension	0	0	0	1
Myocardial infarction	0	0	1	0
Cancer	4	3	2	1
Heart failure	4	6	2	1
Renal failure	2	0	0	0
Complications from lung transplantation	1	0	0	0
Complications from hip fracture	0	0	1	0
Gastrointestinal tract failure	1	1	0	0
Sepsis	1	0	0	0
Seizures	0	1	0	0
Infection	0	1	0	0
Other	1	0	0	0
Unknown	19	11	2	4
Organ failure				
No organ failure	63 (80%)	60 (76%)	58 (79%)	58 (84%)
Any organ failure	14 (18%)	19 (24%)	8 (11%)	7 (10%)
No data available	2 (3%)	0	7 (10%)	4 (6%)
Type of organ failure†				
Supplementary oxygen use	12	17	5	6
Lung transplant§	1	2	0	1
Dialysis	2	0	0	1
Total parenteral nutrition	0	0	1	1
Cardiac ablation	2	1	1	2
Pacemaker	0	1	1	0
Malignancy	7	6	0	1
Development of pulmonary hypertension	Not collected	Not collected	2	9
Median follow-up time (IQR), months	97.4 (37.3–121.9)	86.7 (32.3–120.2)	40.78 (24.9–54.5)	42.2 (36.4–58.7)

CYC, cyclophosphamide; ILD, interstitial lung disease; MMF, mycophenolate ; SLS, Scleroderma Lung Studies; SSc, systemic sclerosis.

*These patients were not in death registries, but were unreachable.

†Subjects could have more than one cause of death and type of organ failure recorded. In SLS I CYC group, 1 person died of sepsis and respiratory failure, 1 died of heart and renal failure, 1 died of respiratory failure and renal failure, and 1 died of respiratory failure and gastrointestinal failure. In SLS I placebo group, 1 person died of respiratory failure and heart failure, 1 died of respiratory failure and gastrointestinal failure, 1 died of respiratory failure and infection, and 1 died of respiratory failure and seizures.

‡In the SLS I study, 2 deaths (1 CYC, 1 placebo) due to 'Respiratory Failure' were not attributed to underlying SSc based on the case report form.

§No patients received heart, kidney, liver or bone marrow transplantations during the follow-up period.

I included the anus (n=1), colon (n=2), vulvar (n=1), prostate (n=1), sarcoma (n=1) and breast (n=1) within the CYC arm, and colon (n=1), oesophageal (n=1), lung (n=3) and Hodgkin's lymphoma (n=1) within the placebo arm.

There is no difference in long-term survival between patients randomised to MMF versus CYC in SLS II

Over twice as many deaths occurred in the CYC arm (n=11) compared with the MMF arm (n=5) during the 24-month study period in SLS II (p=0.160 by log-rank test).[7] During the 8-year long-term follow-up period, an additional five deaths occurred in the CYC arm compared with an additional nine deaths in the MMF arm. There was no significant difference in the time to death (p=0.627 by log-rank test; figure 2A) nor the time to death or organ failure (p=0.343 by log-rank test; figure 2B) for patients randomised to CYC versus MMF in SLS II. However, there appeared to be a separation in the survival curves favouring MMF within the first 2 years (figure 2A). There was no significant difference in the time to the development of organ failure between the two groups (p=0.692 by log-rank test; online supplementary figure S3). Two malignancies occurred during the follow-up period (MMF: n=1 thyroid cancer, n=1 papillary urothelial carcinoma; CYC: none).
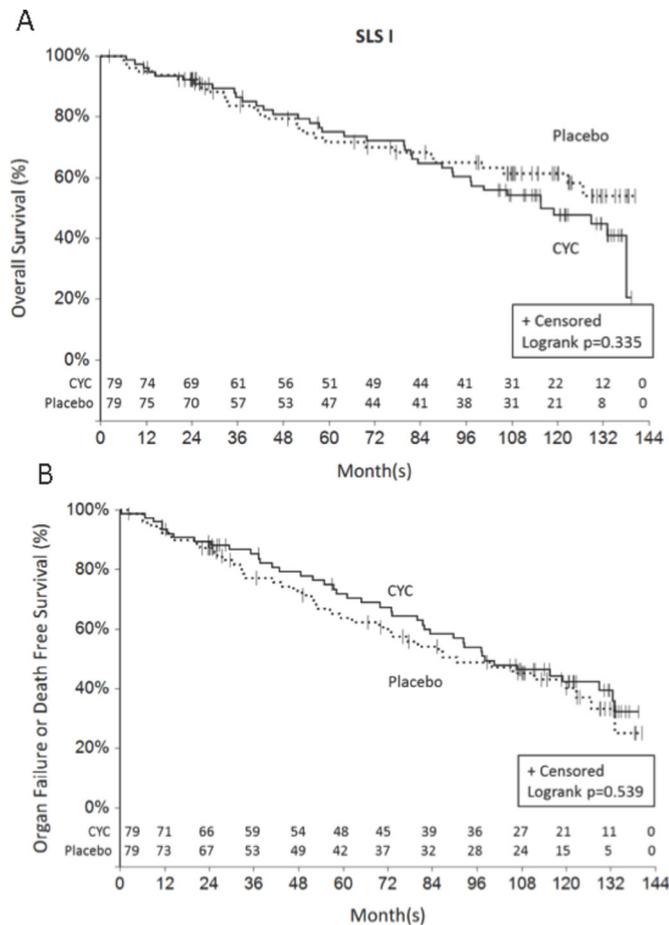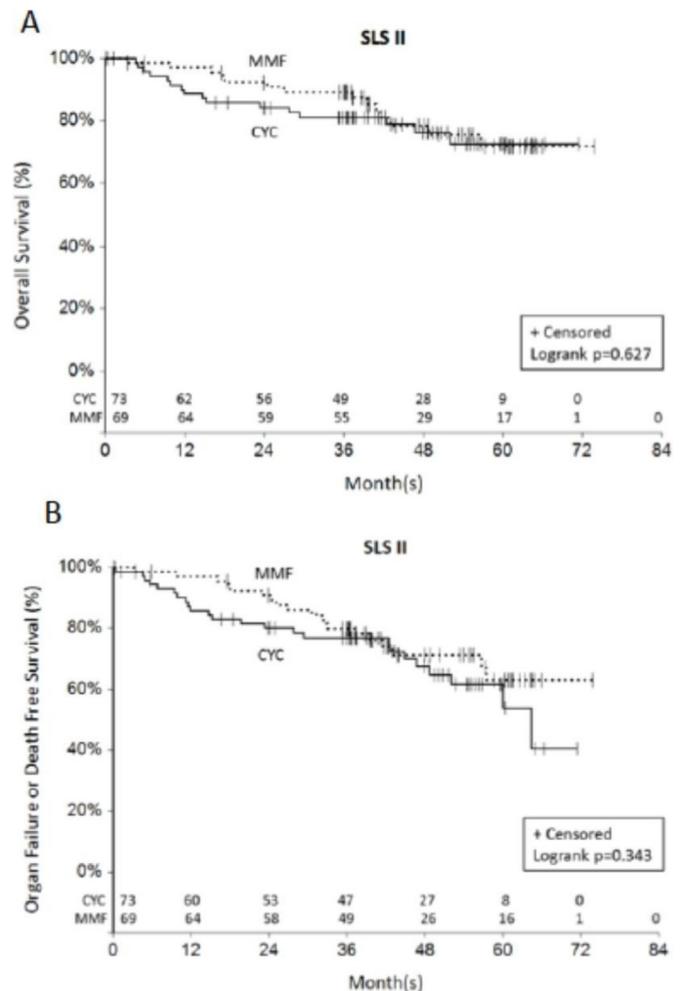
Figure 1 Time to death (A) and time to death or organ failure (B) from randomisation in SLS I. There was no significant difference in the time to death (p=0.335 by log-rank test; A) nor the time to death or organ failure (p=0. 539 by log-rank test; B) for patients randomised to CYC versus placebo in SLS I. The last known date they were known to be alive was used for the survival analysis. CYC, cyclophosphamide; SLS, Scleroderma Lung Studies.

Figure 2 Time to death (A) and time to death or organ failure (B) from randomisation in SLS II. There was no significant difference in the time to death (p=0.627 by log-rank test; (A)) nor the time to death or organ failure (p=0.343 by log-rank test; (B)) for patients randomised to CYC versus MMF in SLS II. The last known date they were known to be alive was used for the survival analysis. CYC, cyclophosphamide; MMF, mycophenolate mofetil; SLS, Scleroderma Lung Studies.

Longitudinal assessments of FVC and DLCO predict long-term survival in SLS I and II

SLS I: Cox proportional hazards models and joint models

The basic model developed from the SLS I cohort, as described in the online supplementary methods, consisted of the following covariates: treatment arm (CYC vs placebo), baseline extent of cutaneous sclerosis (MRSS), age at randomisation (years) and sex. Among these variables, increased age and MRSS were associated with increased mortality. The following variables were independently associated with mortality in the final models: (1) baseline FVC%-predicted; (2) longitudinal assessment of the FVC%-predicted measured as a time-varying covariate over 24 months; and (3) longitudinal assessment of the DLCO%-predicted measured as a time-varying covariate over 24 months (table 3). None of the quantitative lung fibrosis/ILD scores was associated with long-term survival when added to the base model. The Akaike information criterion (AIC),[11–13] which estimates the quality of each model relative to each of the other models, was slightly lower (better) for the models that included the longitudinally measured FVC and DLCO parameters compared with those that included the baseline FVC and DLCO parameters (table 3). Thus, the final SLS I survival models demonstrated that decreased age, decreased extent of cutaneous sclerosis, as well as an improved course of the

FVC%-predicted and the DLCO%-predicted over 24 months were associated with better survival outcomes (table 3). As stated in the online supplementary methods, we created separate models for the FVC and DLCO variables to avoid collinearity.

SLS I: joint model validation analysis

Using the same basic model as above (eg, MRSS, age), the longitudinal assessment of the FVC%-predicted was significantly associated with the outcome (table 3). When added to the basic model, the longitudinal assessment of the DLCO%-predicted was also significantly associated with the outcome (table 3). As noted above, the programming for the joint model does not allow for the inclusion of two longitudinally measured covariates simultaneously.

SLS I: exploratory analyses

In an exploratory analysis, we examined whether the change from baseline in the FVC%-predicted and DLCO%-predicted at 12 months predicted survival. When added to the basic model (eg, MRSS, age), none of the change scores was significantly associated with survival. In addition, we explored whether combined, categorical changes in the FVC%-predicted and DLCO%-predicted

Table 3 Final models for predicting death in SLS I

	Cox model using baseline FVC as covariate	Cox model using FVC as time-dependent covariate	Joint model using FVC as a time-dependent covariate
	HR (95% CI)	HR (95% CI)	HR (95% CI)
FVC models			
Baseline MRSS	1.03 (1.01 to 1.05)*	1.03 (1.01 to 1.06)**	1.03 (1.01 to 1.06)**
Age (years)	1.06 (1.04 to 1.08)†	1.06 (1.04 to 1.08)†	1.06 (1.03 to 1.08)†
FVC%-predicted	0.97 (0.95 to 0.99)**	0.96 (0.94 to 0.97)†	0.97 (0.96 to 0.98)†
AIC	564	546	789
	Cox model using baseline DLCO as covariate	Cox model using DLCO as time-dependent covariate	Joint model using DLCO as a time-dependent covariate
	HR (95% CI)	HR (95% CI)	HR (95% CI)
DLCO models			
Baseline MRSS	1.04 (1.01 to 1.06)**	1.04 (1.02 to 1.07)**	1.04 (1.01 to 1.06)**
Age (years)	1.05 (1.03 to 1.08)†	1.05 (1.02 to 1.07)†	1.05 (1.03 to 1.07)†
DLCO%-predicted	0.98 (0.96 to 1.00)	0.96 (0.95 to 0.98)†	0.96 (0.94 to 0.98)†
AIC	569	525	791

AIC, Akaike information criterion; FVC, forced vital capacity; DLCO, single-breath diffusing capacity for carbon monoxide; MRSS, Modified Rodnan Skin Score; SLS, Scleroderma Lung Studies.
*P<0.05, **P<0.01, ***P<0.001.
†P<0.0001.

at 12 and 24 months predicted survival. None of the categorical declines at 12 months (eg, FVC decline ≥10%; FVC decline ≥15%; DLCO decline ≥15%; FVC decline ≥10% and DLCO decline ≥15%; FVC decline ≥10% or DLCO decline ≥15%) was significantly associated with long-term survival when added to the basic model; however, these individual categorical declines at 24 months were associated with long-term survival when added to the basic model (online supplementary table S1). Too few (patient (n=1) experienced an FVC decline 5%–9% and DLCO decline ≥15% to include this covariate in the model).

SLS II: Cox proportional hazards model
The basic model developed from the SLS II cohort consisted of the following covariates: treatment arm (CYC vs MMF), baseline extent of cutaneous sclerosis (MRSS), age at randomisation

(years) and sex. Similar to SLS I, increased age and increased MRSS at baseline were associated with increased mortality in SLS II. Baseline FVC%-predicted and baseline DLCO%-predicted were not significantly associated with time to death when added to the basic model that comprised age and MRSS. However, the longitudinal assessment of the FVC%-predicted and the longitudinal assessment of the DLCO%-predicted were each associated with the outcome when added to the basic model (table 4). None of the quantitative lung fibrosis/ILD scores was associated with long-term survival when added to the base model. Therefore, the final SLS II survival models demonstrated that decreased age, decreased extent of cutaneous sclerosis, as well as an improved course of the FVC%-predicted and the DLCO%-predicted over 24 months were associated with better survival outcomes (table 4).

Table 4 Final models for predicting death in SLS II

	Cox model using baseline FVC as covariate	Cox model using FVC as time-dependent covariate	Joint model using FVC as a time-dependent covariate
	HR (95% CI)	HR (95% CI)	HR (95% CI)
FVC models			
Baseline MRSS	1.49 (0.82 to 2.56)	1.61 (0.93 to 2.77)	1.66 (0.98 to 2.83)
Age (years)	1.09 (1.04 to 1.14)***	1.09 (1.04 to 1.14)***	1.08 (1.04 to 1.13)***
FVC%-predicted	1.17 (0.64 to 2.14)	0.48 (0.28 to 0.81)**	0.51 (0.30 to 0.88)*
AIC	250	217	347
	Cox model using baseline DLCO as covariate	Cox model using DLCO as time-dependent covariate	Joint model using DLCO as a time-dependent covariate
	HR (95% CI)	HR (95% CI)	HR (95% CI)
DLCO models			
Baseline MRSS	1.56 (0.92 to 2.67)	1.86 (1.00 to 3.47)	1.85 (1.07 to 3.19)*
Age (years)	1.08 (1.03 to 1.13)***	1.08 (1.02 to 1.15)**	1.07 (1.03 to 1.12)**
DLCO%-predicted	0.95 (0.85 to 1.06)	0.83 (0.72 to 0.96)*	0.84 (0.76 to 0.94)**
AIC	249	150	360

AIC, Akaike information criterion; DLCO, single-breath diffusing capacity for carbon monoxide; FVC, forced vital capacity; MRSS, Modified Rodnan Skin Score.
*P<0.01, **P<0.01, ***P<0.001.
†P<0.0001.

SLS II: joint model validation analysis

Using the same basic model as above (eg, MRSS, age), the longitudinal assessment of the FVC%-predicted was significantly associated with the outcome (table 4). When added to the basic model, the longitudinal assessment of the DLCO%-predicted was also significantly associated with the outcome (table 4).

SLS II: exploratory analyses

In an exploratory analysis, we examined whether the change from baseline in the FVC%-predicted and DLCO%-predicted at 12 months predicted survival. When added to the basic model (eg, MRSS, age), the change in FVC from baseline to 12 months predicted long-term survival (estimate 0.52 (CI 0.31 to 0.90); p<0.05), but not the change in DLCO from baseline to 12 months. In addition, we explored whether combined, categorical changes in the FVC%-predicted and DLCO%-predicted at 12 months predicted survival as described above. We found that an FVC decline ≥10% at 12 months and FVC decline ≥15% at 12 months were each associated with long-term survival when added to the base model with the covariates of MRSS and age (online supplementary table S2). Patients in SLS II had a decline in DLCO ≥15% at 12 months in SLS II; therefore, we were unable to analyse any of the the composite categorical decline variables that included the DLCO.

Use of potential disease-modifying agents

While no data were available regarding the use of potential disease-modifying agents beyond the 24-month study period in SLS I, in SLS II these data were collected and demonstrated that the majority of patients consumed MMF following the 24-month study period. (Please see online supplementary table S3 for a list of all potential disease-modifying agents used following cessation of study drug in SLS II.)

DISCUSSION

The present study is the first study to examine mortality outcomes in patients who participated in two of the largest RCTs for SSc-ILD. The results presented herein demonstrate that treatment with 1 year of CYC compared with placebo does not improve long-term survival outcomes in patients with SSc-ILD. The findings also demonstrate that there is no difference in long-term survival between patients randomised to CYC versus MMF in SLS II.

Both SLS I and II demonstrated that treatment with immunosuppression led to short-term improvements in surrogate measures of SSc-ILD outcomes.[6 7] However, the present findings suggest that short-term treatment with CYC and MMF may not improve long-term outcomes of patients with SSc-ILD. Where known, the majority of patients in both SLS I and II died of complications related to SSc, and respiratory failure due to end-stage lung disease was the leading cause of death. These findings are consistent with recent reports of mortality outcomes in observational cohorts.[1 2]

Relatively few malignancies occurred in either cohort. Furthermore, although studies have demonstrated an increase in haematological malignancies in SSc,[14] only one case of lymphoma occurred in the placebo arm of SLS I. The paucity of malignancies observed in both cohorts may in part be due to the length of the follow-up period, especially in SLS II where the median follow-up was only 4 years, as well as the observation that respiratory failure was the leading cause of death.

These findings highlight a need to determine the appropriate duration of treatment for SSc-ILD. In SLS I, very few patients reported use of immunosuppression in year 2 of the study despite the accepted view that ILD progression generally occurs up to 5 years from disease onset in SSc. This may explain why there was no difference in long-term survival between the two study arms in SLS I. It is plausible that initiation of a maintenance therapy regimen after induction therapy may affect long-term survival outcomes. In SLS II, many patients continued on immunosuppression after the trial concluded (MMF was most common). The use of MMF in the CYC arm may in part explain why the trend for an MMF-related survival benefit observed in the first 2 years diminished in the subsequent follow-up years.

More research is needed to determine the appropriate length of treatment for immunosuppression in SSc-ILD. A prior small retrospective study demonstrated improvement/stability in the FVC after 18 months of azathioprine (AZA) maintenance therapy in patients who first completed a 6-month induction course of intravenous CYC for SSc-ILD.[15] However, without a control arm, it is impossible to discern whether the observed improvement/stability in the FVC represents a true treatment response versus the natural course of ILD in SSc. An additional study demonstrated that patients who responded favourably to pulse CYC and were subsequently treated with AZA experienced a higher rate of improvement or stabilisation in lung function compared with patients who did not respond to pulse CYC.[16]

The present analysis also revealed significant predictors of long-term mortality in SSc-ILD. In line with prior observational studies,[17–19] increased skin score and increased age were independently associated with increased mortality. In contrast to prior observational studies,[18 20–22] however, male gender and African–American race were not associated with an increased risk of mortality. Regarding gender, the SLS I and II cohorts comprised predominantly of women; thus, these studies may be underpowered to detect true gender differences in long-term survival. In terms of race, our findings could potentially suggest that in the context of a clinical trial, in which all patients have equal access to healthcare and uniform follow-up, race does not play a substantial role in predicting long-term survival.

Consistent with prior observational studies,[16 19 23 24] low baseline FVC was associated with an increased risk of mortality. However, the course of the FVC and the DLCO over 24 months appeared to be more robust predictors of long-term survival in both SLS I and II than the baseline measurements of these parameters when comparing the AIC for the models. The individual parameter estimates were similar for both the Cox model and the joint model that we used as a validation approach, suggesting that the relationship between survival and FVC (or DLCO) is not biased by non-ignorable missing data.

A recent single-centre observational cohort study of patients with SSc-ILD also found that pulmonary function trends at 1 and 2 years predicted intermediate to long-term mortality.[25] This study demonstrated that 1-year categorical trends in the FVC and DLCO were the most accurate prognostic determinants of mortality, while at 2 years changes in gas transfer were the most important predictors of mortality.[25] As this was an observational study, the authors could not adequately control for treatment effect and selection bias. We were able to replicate some, but not all of these findings in the SLS I or II cohorts. Taken together, the findings of the present study provide further evidence that trends in pulmonary function may offer more prognostic information than baseline pulmonary function measurements. This may in part be due to the fact that substantial variability exists in a single FVC and DLCO measurement. Repeated measurements of the FVC/DLCO may yield more clinically meaningful information regarding ILD progression and survival.

The findings of the present study should be interpreted in the context of specific limitations. There were subtle differences in the baseline characteristics of the SLS I and SLS II cohorts. For instance, the DLCO was lower in SLS I, and this may have been due to less scrupulousness in excluding pulmonary arterial hypertension (PAH). The Baseline Dyspnoea Index (BDI) score was also lower in SLS I, although these differences could be related to using different instruments to administer the BDI in the two studies.[29] The quantitative extent of interstitial lung disease/quantitative extent of lung fibrosis in the zone of maximal involvement on HRCT was higher in SLS II, while the quantitative extent of interstitial lung disease in whole lung on HRCT was higher in SLS I. However, in SLS I, non-volumetric CT scans of 1–2 mm slice thickness were acquired at 10 mm increments, while in SLS II volumetric CT scans of 1–1.5 mm slice thickness were acquired contiguously. Overall, the two cohorts were strikingly similar.

A number of SLS I and, to a lesser degree, SLS II participants were lost to follow-up during the course of this long-term follow-up study. This can introduce bias, especially in cases where early censoring occurred. The Cox proportional hazards model was used to deal with time to event data in the presence of censoring. Moreover, while a morbidity and mortality committee adjudicated the causes of death during the 24-month study periods, less detailed information was obtained during long-term follow-up periods regarding causes of death. In addition, although we successfully collected information on the use of immunosuppression following the conclusion of SLS II, the timing, duration and dosages of immunosuppression reported by patients varied widely and precluded any kind of meaningful statistical analysis of these data (in SLS I, very few patients continued on immunosuppression during the second year; beyond this point no data were available regarding immunosuppression use). Finally, while PH was identified as the cause of death in only one of the SLS II patients, this comorbidity may have influenced survival rates in both cohorts.

Notable strengths of the present manuscript include the use of two relatively large, well-characterised SSc-ILD cohorts undergoing standard treatment approaches with uniform follow-up measurements over the course of 2 years. These cohorts comprised patients from multiple SSc Centers of Excellence across the USA, augmenting the generalisability of our study findings to a US population. Furthermore, we identified the same mortality predictor variables in both cohorts, suggesting that our results are likely reproducible in other similar SSc cohorts. Finally, we used a joint model as a means of internal validation.

In summary, the findings of the present analyses demonstrate that increased baseline skin score, increased baseline age, and the course of the FVC and DLCO over 2 years are important predictors of long-term survival in SSc-ILD. Treatment with immunosuppression may not improve long-term survival in patients with SSc-ILD, in contrast to haematopoetic stem cell transplantation,[26–28] which seems to offer a more sustained improvement in long-term survival and may especially help those patients who have early, rapidly progressive SSc with organ involvement. Future studies are needed to determine how the duration of immunosuppression affects long-term survival among patients with SSc-ILD. With the emergence of promising new therapies for SSc-ILD (eg, antifibrotics, or combination therapy with antifibrotics and immunosuppression), additional studies are needed to compare how these novel approaches affect survival compared with the current standard of care for SSc-ILD.

Author affiliations
[1]Department of Medicine, University of California, Los Angeles, David Geffen School of Medicine, Los Angeles, California, USA
[2]Department of Biomathematics, University of California, Los Angeles, California, USA
[3]NIAID, NIH, Bethesda, Maryland, USA
[4]Rho Federal Systems, Inc., Chapel Hill, North Carolina, USA
[5]Department of Medicine, University of Michigan Medical School, Ann Arbor, Michigan, USA
[6]Department of Medicine, Georgetown University, Washington, District of Columbia, USA
[7]Department of Medicine, University of Illinois Hospital, Chicago, Illinois, USA
[8]Department of Medicine, Robert Wood Johnson Medical School, New Brunswick, New Jersey, USA
[9]Department of Medicine, Duke University, Durham, North Carolina, USA

REFERENCES

1 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* 2012;51:1017–26.

2 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;69:1809–15.

3 Wells AU. Interstitial lung disease in systemic sclerosis. *La Presse Médicale* 2014;43:e329–43.

4 D'Angelo WA, Fries JF, Masi AT, *et al*. Pathologic observations in systemic sclerosis (scleroderma). A study of fifty-eight autopsy cases and fifty-eight matched controls. *Am J Med* 1969;46:428–40.

5 Volkmann ER, Tashkin DP. Treatment of systemic sclerosis-related interstitial lung disease: a review of existing and emerging therapies. *Ann Am Thorac Soc* 2016;13:2045–56.

6 Tashkin DP, Elashoff R, Clements PJ, *et al*. Cyclophosphamide versus placebo in scleroderma lung disease. *N Engl J Med* 2006;354:2655–66.

7 Tashkin DP, Roth MD, Clements PJ, *et al*. Mycophenolate mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease (SLS II): a randomised controlled, double-blind, parallel group trial. *Lancet Respir Med* 2016;4:708–19.

8 Volkmann ER, Chung A, Tashkin DP. Managing systemic sclerosis-related interstitial lung disease in the modern treatment era. *J Scleroderma Relat Disord* 2017;2:72–83.

9 Zhang D, Chen MH, Ibrahim JG, *et al*. JMFit: A SAS macro for joint models of longitudinal and survival data. *J Stat Softw* 2016;71.

10 Tashkin DP, Elashoff R, Clements PJ, *et al*. Effects of 1-year treatment with cyclophosphamide on outcomes at 2 years in scleroderma lung disease. *Am J Respir Crit Care Med* 2007;176:1026–34.

11 Bozdogan H. Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions. *Psychometrika* 1987;52:345–70.

12 Snipes M, Taylor DC. Model selection and Akaike Information Criteria: an example from wine ratings and prices. *Wine Economics and Policy* 2014;3:3–9.

13 Symonds MRE, Moussalli A. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behav Ecol Sociobiol* 2011;65:13–21.

14 Zeineddine N, Khoury LE, Mosak J. Systemic sclerosis and malignancy: a review of current data. *J Clin Med Res* 2016;8:625–32.

15 Bérezné A, Ranque B, Valeyre D, *et al*. Therapeutic strategy combining intravenous cyclophosphamide followed by oral azathioprine to treat worsening interstitial lung disease associated with systemic sclerosis: a retrospective multicenter open-label study. *J Rheumatol* 2008;35:1064–72.

16 Iudici M, Cuomo G, Vettori S, *et al*. Low-dose pulse cyclophosphamide in interstitial lung disease associated with systemic sclerosis (SSc-ILD): efficacy of maintenance immunosuppression in responders and non-responders. *Semin Arthritis Rheum* 2015;44:437–44.

17 Steen VD, Conte C, Owens GR, *et al*. Severe restrictive lung disease in systemic sclerosis. *Arthritis Rheum* 1994;37:1283–9.

18 Mayes MD, Lacey JV, Beebe-Dimmer J, *et al*. Prevalence, incidence, survival, and disease characteristics of systemic sclerosis in a large US population. *Arthritis Rheum* 2003;48:2246–55.

19 Nihtyanova SI, Schreiber BE, Ong VH, *et al*. Prediction of pulmonary complications and long-term survival in systemic sclerosis. *Arthritis Rheumatol* 2014;66:1625–35.

20 Goldin J, Elashoff R, Kim HJ, *et al*. Treatment of scleroderma-interstitial lung disease with cyclophosphamide is associated with less progressive fibrosis on serial thoracic high-resolution CT scan than placebo: findings from the scleroderma lung study. *Chest* 2009;136:1333–40.

21 Domsic RT, Nihtyanova SI, Wisniewski SR, *et al*. Derivation and external validation of a prediction rule for five-year mortality in patients with early diffuse cutaneous systemic sclerosis. *Arthritis Rheumatol* 2016;68:993–1003.

22 Mendoza F, Derk CT. Systemic sclerosis mortality in the United States: 1999-2002 implications for patient care. *J Clin Rheumatol* 2007;13:187–92.

23 Assassi S, Sharif R, Lasky RE, *et al*. Predictors of interstitial lung disease in early systemic sclerosis: a prospective longitudinal study of the GENISOS cohort. *Arthritis Res Ther* 2010;12:R166.

24 Morgan C, Knight C, Lunt M, *et al*. Predictors of end stage lung disease in a cohort of patients with scleroderma. *Ann Rheum Dis* 2003;62:146–50.

25 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 Rheumatol* 2017;69:1670–8.

26 Sullivan KM, Goldmuntz EA, Keyes-Elstein L, *et al*. Myeloablative autologous stem-cell transplantation for severe scleroderma. *N Engl J Med* 2018;378:35–47.

27 van Laar JM, Farge D, Sont JK, *et al*. Autologous hematopoietic stem cell transplantation vs intravenous pulse cyclophosphamide in diffuse cutaneous systemic sclerosis: a randomized clinical trial. *JAMA* 2014;311:2490–8.

28 Shouval R, Furie N, Raanani P, *et al*. Autologous hematopoietic stem cell transplantation for systemic sclerosis: a systematic review and meta-analysis. *Biol Blood Marrow Transplant* 2018;24:937–44.

29 Mahler DA, Weinberg DH, Wells CK, *et al*. The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest* 1984;85:751–8.

30 Mahler DA, Ward J, Fierro-Carrion G, *et al*. Development of self-administered versions of modified baseline and transition dyspnea indexes in COPD. *COPD* 2004;1:165–72.

32 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.