

Published in final edited form as:

Thorax. 2009 July ; 64(7): 626–630. doi:10.1136/thx.2008.111393.

Exercise peripheral oxygen saturation (SpO₂) accurately reflects arterial oxygen saturation (SaO₂) and predicts mortality in systemic sclerosis

J J Swigris¹, X Zhou¹, F S Wamboldt², R du Bois¹, R Keith¹, A Fischer¹, G P Cosgrove¹, S K Frankel¹, D Curran-Everett³, and K K Brown¹

¹ Interstitial Lung Disease Program and Autoimmune Lung Center, National Jewish Health, Denver, Colorado, USA

² Division of Psychosocial Medicine, National Jewish Health, Denver, Colorado, USA

³ Division of Biostatistics and Bioinformatics, National Jewish Health, Denver, Colorado, USA

Abstract

Background—Measures of oxygenation have not been assessed for prognostic significance in systemic sclerosis-related interstitial lung disease (SSc-ILD).

Methods—83 subjects with SSc-ILD performed a maximal cardiopulmonary exercise test with an arterial line. The agreement between peripheral oxygen saturation (SpO₂) and arterial oxygen saturation (SaO₂) was examined and survival differences between subgroups of subjects stratified on SpO₂ were analysed. Cox proportional hazards analyses were used to examine the prognostic capabilities of SpO₂.

Results—At maximal exercise the mean (SD) difference between SpO₂ and SaO₂ was 2.98 (2.98) and only 15 subjects had a difference of >4 points. The survival of subjects with SSc-ILD whose maximum exercise SpO₂ (SpO₂max) fell below 89% or whose SpO₂max fell >4 points from baseline was worse than subjects in comparator groups (log rank p = 0.01 and 0.01, respectively). The hazard of death during the median 7.1 years of follow-up was 2.4 times greater for subjects whose SpO₂max fell below 89% (hazard ratio 2.4, 95% CI 1.1 to 4.9, p = 0.02) or whose SpO₂max fell >4 points from baseline (hazard ratio 2.4, 95% CI 1.1 to 5.0, p = 0.02).

Conclusion—In patients with SSc-ILD, SpO₂ is an adequate reflection of SaO₂ and radial arterial lines need not be inserted during cardiopulmonary exercise tests in these patients. Given the ease of measurement and its prognostic value, SpO₂ should be considered as a meaningful clinical and research outcome in patients with SSc-ILD.

Measurement of static pulmonary physiology^{1–4} and, more recently, assessments of high-resolution CT (HRCT) scans⁵ have been shown to provide important prognostic information about patients with fibrosing interstitial lung disease related to systemic sclerosis (SSc-ILD). For patients with idiopathic fibrosing ILD, assessments of blood oxygenation—particularly those done while patients are exerting—have also been shown to predict outcome.^{6–8}

Measures of peripheral oxygen saturation (SpO₂) are used to estimate arterial blood saturation (SaO₂); however, the validity of the relationship between SpO₂ and SaO₂ is dependent on a number of factors including the adequacy of peripheral perfusion.⁹ Given

Correspondence to: Dr J J Swigris, Interstitial Lung Disease Program and Autoimmune Lung Center, National Jewish Medical and Research Center, 1400 Jackson Street, Denver, CO 80206, USA; swigrisj@njc.org.

Competing interests: None.

recent data questioning the reproducibility of the level of desaturation (as measured by SpO₂) over short time intervals in patients with idiopathic pulmonary fibrosis (IPF)¹⁰ but continued enthusiasm to use this easily obtainable measure in patients with fibrosing ILD, we sought to investigate SpO₂ in patients with SSc-ILD—a disease in which peripheral circulation is often impaired, an impairment considered to preclude reliable interpretation of peripheral oxygen saturation assessments. We performed this study to test two hypotheses: (1) SpO₂ would inaccurately reflect SaO₂ at rest and the inaccuracy would be even greater during maximal exercise; and (2) despite the hypothesised inaccuracy, peripheral measures of blood oxygenation during exertion should provide significant prognostic information in patients with SSc-ILD. Thus, besides assessing the usefulness of SpO₂ as a prognostic marker, a goal of the study was to determine whether an arterial line is needed to accurately assess exertional oxygenation in patients with SSc-ILD.

METHODS

Subjects

We identified 83 patients with SSc-ILD and no signs of pulmonary hypertension on physical examination who were evaluated in the ILD Program at National Jewish Health with pulmonary function tests and a maximal cardiopulmonary test between 1983 and 2005.

All subjects were evaluated clinically using a standard protocol focused on the identification of the cause of ILD. Patients with SSc met the diagnostic criteria adopted by the American College of Rheumatology;¹¹ those with SSc sine scleroderma (ssSSc) met the criteria suggested by Poormoghim and colleagues.¹² Patients with overlap syndromes were excluded. In patients with SSc the diagnosis of ILD was made by surgical biopsy (n = 17), chest radiography (n = 60) or CT scanning (n = 4).

Maximal cardiopulmonary exercise test

All subjects underwent maximal cardiopulmonary exercise testing at our institution according to a standardised protocol. A radial arterial line and a peripheral pulse oximeter were placed before commencing exercise. The peripheral pulse oximeter was placed on the index finger of the hand opposite the arterial line. If an adequate pulse oximeter signal was not obtained (inability to obtain correct pulse), an earlobe probe was used. In more than 95% of subjects, finger probes were used. Baseline measurements were collected after the patient mounted the cycle ergometer (Vmax 29; SensorMedics, Yorba Linda, California, USA) and just before beginning to pedal. Subjects pedalled for 3 min at 60 rpm and then work was incrementally added every minute; the goal was to reach a subject's maximal exercise capacity within 6–12 min. Blood was drawn from the arterial line at baseline and after every minute of exercise, and analysed with a co-oximeter (to measure the percentage of arterial oxygen saturation (SaO₂)) and blood gas analyser.

Statistical analysis

Categorical data are presented as counts or percentages. Continuous data are presented as mean values with standard deviation (SD) or median values with interquartile range (IQR). Graphical methods, including Bland-Altman plots,¹³ were used to display agreement between SpO₂ and SaO₂; for each subject the *average* of the SpO₂ and SaO₂ at maximal exercise was plotted against the *difference* between the SpO₂ and SaO₂ at maximal exercise. In certain analyses we used a difference between the SpO₂ and SaO₂ of >4 percentage points as the variable of interest. Four percentage points was chosen as the difference value of significance because this difference falls outside the intrinsic variability of the pulse oximeters used and is felt to be clinically relevant.⁷⁸ The product-limit method was used to estimate survival probabilities and the Kaplan–Meier method was used to generate survival

curves which were compared using the log-rank test. Cox proportional hazard models were used to examine the prognostic capabilities of SpO₂. We confirmed the proportionality assumption was met for the dichotomised SpO₂ variables by examining log(-log) plots. All statistical analyses were performed using SAS Version 9.1 (SAS Institute, Cary, North Carolina, USA). A p value of <0.05 was considered to be statistically significant.

RESULTS

The baseline characteristics of the subjects are shown in table 1. SpO₂ overestimated SaO₂ both at rest and at maximal exercise. The mean (SD) difference between SpO₂ and SaO₂ (keeping overestimates as positive values and underestimates as negative values) was similar at maximal exercise and rest (1.5 (4) vs 1.5 (3), p = 1, fig 1). Similar results were obtained when absolute values of the difference between SpO₂ and SaO₂ were used (2.24 (1.91) vs 2.98 (2.98), p = 0.06). SpO₂ misclassified four subjects at maximum exercise (SpO₂ observed to be >88% but SaO₂ was 88%). For these subjects, median (IQR) values for baseline SpO₂, SaO₂ and arterial oxygen tension (PaO₂) and maximum exercise SpO₂, SaO₂ and PaO₂ were 91.5 (89–96), 88.5 (88–90), 61 (59.5–67), and 89.5 (89–93), 82 (79–88), 47.5 (43.5–62.5).

Over the study period (median follow-up 10.3 years, IQR 4–17) there were 40 deaths. Truncating follow-up at 20 years, we observed 39 deaths over the study period (median follow-up after truncation 7.1 years). Median survival was 9.5 years (IQR 4–16). Subjects whose SpO₂ at maximum exercise (SpO₂max) fell to <89% had shorter survival than those whose SpO₂max remained 89% (p = 0.01, log rank test; fig 2). The results were similar when stratifying subjects on whether SpO₂max fell >4 points from baseline (p = 0.01, log rank test). In Cox proportional hazards models, SpO₂ was a significant predictor of mortality; over a median 7.1 years of follow-up, the risk of death was 2.4 times greater for subjects whose SpO₂max fell below 89% (hazards ratio (HR) 2.4, 95% CI 1.2 to 4.9, p = 0.02) than for subjects whose SpO₂max remained 89%. Similarly, the risk of death was 2.4 times greater for subjects whose SpO₂max fell >4 points from baseline (HR 2.4, 95% CI 1.1 to 5.0, p = 0.02) than for subjects whose SpO₂max remained within 4 points of baseline values. When analysed as a continuous variable, the difference between baseline SpO₂ and SpO₂max remained a significant predictor (HR 1.08, 95% CI 1.03 to 1.14, p = 0.002). Controlling for FVC%, the difference between baseline SpO₂ and SpO₂max (continuous variable) remained a significant predictor (HR 1.07, 95% CI 1.01 to 1.14, p = 0.02). Figure 3 shows the relationship between percentage lung carbon monoxide transfer factor (TLCO%) and SpO₂max.

DISCUSSION

With two goals in mind—to assess the usefulness of SpO₂ as a prognostic marker in patients with SSc-ILD and to determine whether an arterial line is needed to accurately assess oxygenation—we conducted a study to first examine agreement between SpO₂ and SaO₂ at rest and maximal exercise and then to analyse the ability of SpO₂ to predict mortality in patients with SSc-ILD. We hypothesised that, in subjects with SSc-ILD, SpO₂ would inaccurately reflect SaO₂ at rest and the disparity would be even greater at maximal exercise. In contrast, we found that SpO₂ was an accurate reflection of SaO₂ both at rest and maximal exertion in these subjects. Moreover, we observed that SpO₂, a simple non-invasive and inexpensive measure to collect, was a predictor of mortality in patients with SSc-ILD.

Recently there has been a groundswell of attention on the use of non-invasive markers of exertional blood oxygenation (eg, nadir SpO₂ during a 6-minute walk test (6MWT) or

statistical manipulations of SpO₂ over the course of a timed walked test) as outcome measures in therapeutic trials and clinical studies enrolling subjects with ILD.^{7,14,15} This increased attention raises three important distinct but related questions regarding the use of SpO₂ at maximal exercise as an outcome metric:

- ▶ Is it valid—does it in fact measure what it is purported to measure (eg, true blood oxygenation or SaO₂)?
- ▶ Is it reliable—if it is measured at two separate time points in a subject whose clinical status has not changed, will it produce similar results?
- ▶ Is it responsive to underlying change—if a subject's blood oxygenation at maximal exercise changes from baseline, will SpO₂ reflect those changes?

The current study shows that SpO₂ at maximal exercise is a valid measure of blood oxygenation at maximal exercise in patients with SSc-ILD, and SpO₂ does in fact accurately track changes in SaO₂.

The Bland-Altman plots reinforce this finding. These plots give a graphical presentation of the agreement between two methods of measurement; they depict an estimate of the bias (or systematic error which is simply the overestimation or underestimation of one measure compared with the other) as the mean difference between the two measures. The precision of that estimate is reflected in its standard deviation. Whereas correlation coefficients express the relationship between two variables, Bland-Altman plots depict agreement between them. When one is trying to determine the accuracy with which one measure (eg, SpO₂) reflects another (eg, SaO₂) or whether one measure might be used in place of another measure, correlation may not tell the true story—there can be extremely high correlation between two measures but, at the same time, poor agreement. This study shows that, for patients with SSc-ILD, SpO₂ is an accurate reflection of SaO₂ at rest or maximal exercise.

Several studies have examined the agreement between SpO₂ and SaO₂⁹ but, to our knowledge, this is the first in a cohort with SSc-ILD. The importance and clinical relevance of this study centres on the peripheral circulation issues in SSc that make most clinicians reluctant to place an arterial line (digits have been lost as a consequence) and wary of SpO₂ accuracy in these patients. There is therefore a need to validate SpO₂ in SSc-ILD. In general, SpO₂ may either overestimate or underestimate SaO₂. In a meta-analysis Jensen and colleagues⁹ reported that, among 23 studies for which bias and precision estimates were available, the absolute mean (SD) bias was 1.99 (0.23) (ie, on average, SpO₂ overestimated SaO₂ by 1.99 points). In those studies the mean (SD) difference between SpO₂ and SaO₂ ranged from -13.2 (8.0) to 12.0 (13.3). The authors commented that severe or rapid desaturation; hypotension, hypothermia, or other unstable haemodynamic or low perfusion states; dysaemoglobinaemia or use of vital dyes; and motion may all confound agreement between SpO₂ and SaO₂. The mean differences between SpO₂ and SaO₂ in the current study fall well within the range mentioned in that analysis.

The results of the current study not only suggest that SpO₂ is a valid surrogate for SaO₂ in patients with SSc-ILD, but also suggest that desaturation, as measured by SpO₂, is a significant predictor of mortality in this patient group. Our results are in line with the work by Lama and colleagues⁸ that suggested desaturation (as measured by SpO₂) during a 6MWT is an important prognosticator in patients with idiopathic interstitial pneumonia. In so far as the 6MWT accurately reflects functional exercise capacity in patients with SSc-ILD—it does so in patients with fibrotic idiopathic interstitial pneumonia¹⁰—we hypothesise that our results would hold for measures of SpO₂ collected during the 6MWT in this patient population. Not surprisingly, we found $\tau_{\text{LCO}}\%$ to be a strong driver of SpO₂max (data not shown); in fact, among several candidate variables including age, gender, FVC% and

baseline SpO₂, TLCO% was the only significant predictor. Like other investigators,¹ we also found TLCO% to be a potent predictor of survival in our cohort (data not shown). Because of the strong relationship between TLCO% and SpO₂max, and because our goal was merely to begin to examine SpO₂max as a prognostic marker, we performed our survival analysis adjusting for FVC% and not TLCO%.

Although the results are novel and clinically relevant, this study has limitations. This is a retrospective analysis of data collected prospectively over a period of three decades. Different pulse and co-oximeters were used during different time periods; however, each instrument is purported to be accurate within two percentage points for SaO₂ values from 70–100% so we can be confident in the readings. Data for this study were collected at a centre situated 5280 feet above sea level. In Denver, patients probably “live” closer to the steep portion of the oxygen dissociation curve (probably on or very close to the shoulder) than patients at lower altitudes. How this affects the results merits consideration and examination in future studies. Given the lack of systematic examinations for pulmonary hypertension and the changes in available technology to assess for pulmonary hypertension over the study period, we cannot be certain how many subjects truly had the condition. Even more complex is the issue of exercise-induced pulmonary hypertension: how many subjects had it is unknown but, as with other studies of subjects with ILD, the possibility of its presence and its effects on exercise SpO₂ must be considered. Given these limitations, we believe the results should be viewed as hypothesis-generating and will hopefully spark continued investigation in this area. These results will need prospective confirmation at other altitudes. Future studies should examine whether SpO₂ values collected during the 6MWT are as meaningful as those collected during cardiopulmonary exercise tests, and efforts should be made to further delineate the relationship between resting or exercise-induced pulmonary hypertension and SpO₂.

CONCLUSION

In SSc-ILD, at both baseline and maximal exercise, SpO₂ is an accurate reflection of SaO₂. In patients with SSc-ILD, SpO₂ carries prognostic value. Because of the ease with which it is assessed, consideration should be given to measuring exercise SpO₂ as a marker of clinical status or as an outcome in clinical trials enrolling subjects with SSc-ILD. Future research could clarify a number of outstanding and important questions related to SpO₂ in this patient population.

REFERENCES

1. Bouros D, Wells AU, Nicholson AG, et al. Histopathologic subsets of fibrosing alveolitis in patients with systemic sclerosis and their relationship to outcome. *Am J Respir Crit Care Med.* 2002; 165:1581–6. [PubMed: 12070056]
2. 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. [PubMed: 12525384]
3. Simeon CP, Armadans L, Fonollosa V, et al. Mortality and prognostic factors in Spanish patients with systemic sclerosis. *Rheumatology (Oxford).* 2003; 42:71–5. [PubMed: 12509616]
4. Steen VD, Medsger TA Jr. Severe organ involvement in systemic sclerosis with diffuse scleroderma. *Arthritis Rheum.* 2000; 43:2437–44. [PubMed: 11083266]
5. 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; 177:1248–54. [PubMed: 18369202]
6. Hallstrand TS, Boitano LJ, Johnson WC, et al. The timed walk test as a measure of severity and survival in idiopathic pulmonary fibrosis. *Eur Respir J.* 2005; 25:96–103. [PubMed: 15640329]

7. Flaherty KR, Andrei AC, Murray S, et al. Idiopathic pulmonary fibrosis: prognostic value of changes in physiology and six-minute-walk test. *Am J Respir Crit Care Med.* 2006; 174:803–9. [PubMed: 16825656]
8. Lama VN, Flaherty KR, Toews GB, et al. Prognostic value of desaturation during a 6-minute walk test in idiopathic interstitial pneumonia. *Am J Respir Crit Care Med.* 2003; 168:1084–90. [PubMed: 12917227]
9. Jensen LA, Onyskiw JE, Prasad NG. Meta-analysis of arterial oxygen saturation monitoring by pulse oximetry in adults. *Heart Lung.* 1998; 27:387–408. [PubMed: 9835670]
10. Eaton T, Young P, Milne D, et al. Six-minute walk, maximal exercise tests: reproducibility in fibrotic interstitial pneumonia. *Am J Respir Crit Care Med.* 2005; 171:1150–7. [PubMed: 15640367]
11. Anon. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum.* 1980; 23:581–90. [PubMed: 7378088]
12. Poormoghim H, Lucas M, Fertig N, et al. Systemic sclerosis sine scleroderma: demographic, clinical, and serologic features and survival in forty-eight patients. *Arthritis Rheum.* 2000; 43:444–51. [PubMed: 10693887]
13. Bland J, Altman D. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999; 8:135–60. [PubMed: 10501650]
14. Azuma A, Nukiwa T, Tsuboi E, et al. Double-blind, placebo-controlled trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med.* 2005; 171:1040–7. [PubMed: 15665326]
15. King TE Jr, Behr J, Brown KK, et al. BUILD-1: a randomized placebo-controlled trial of bosentan in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med.* 2008; 177:75–81. [PubMed: 17901413]

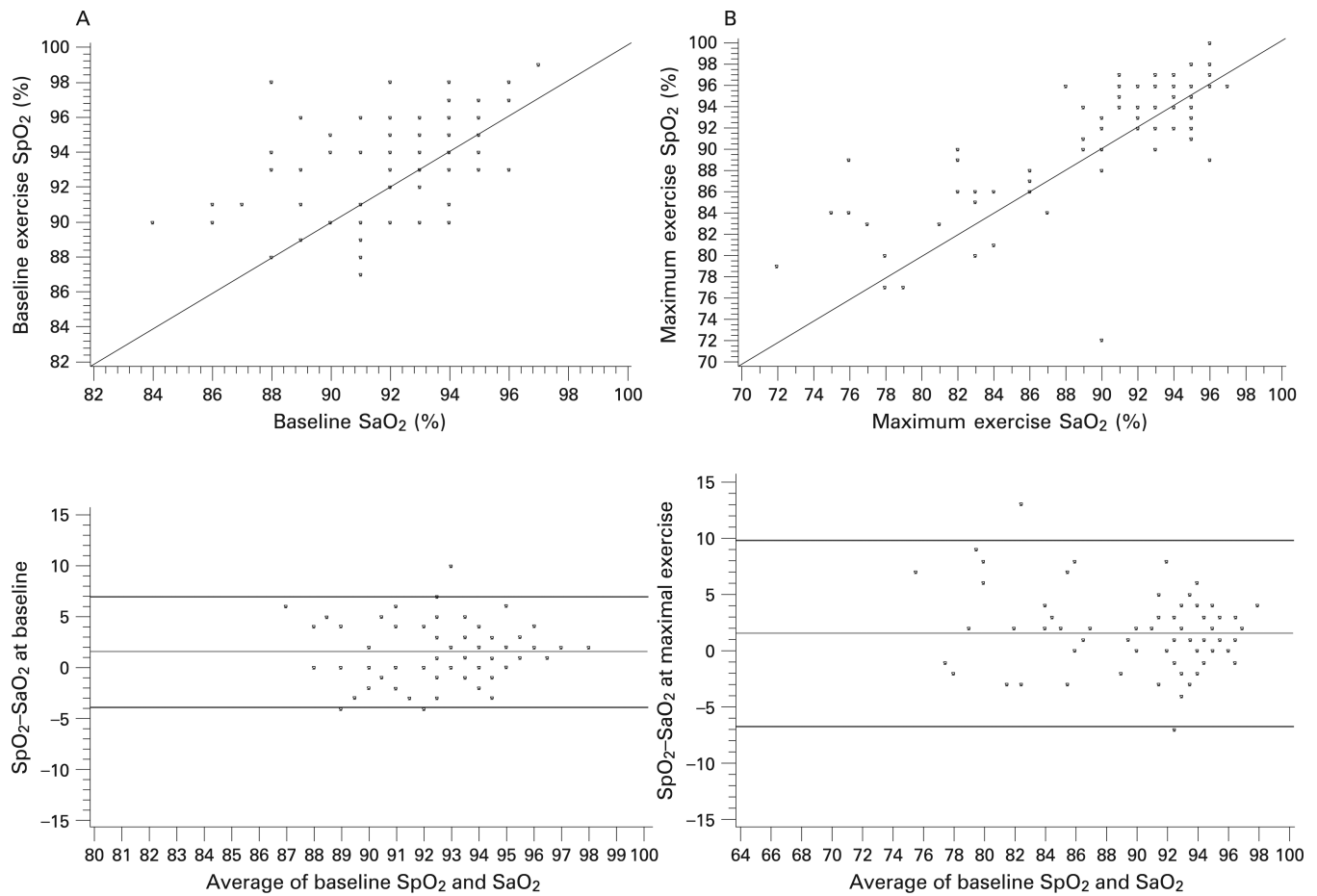


Figure 1.

Plots of peripheral oxygen saturation (SpO₂) against arterial oxygen saturation (SaO₂) and Bland-Altman plots for subjects with fibrosing interstitial lung disease related to systemic sclerosis (SSc-ILD) at (A) baseline and (B) maximum exercise. (A) In the upper graph the diagonal line marks the line of unity, and in the lower graph the middle line marks the mean difference between SpO₂ and SaO₂ and the two outer lines mark twice the standard deviation in either direction. (B) In the upper graph the diagonal line marks the line of unity, and in the lower graph the middle line marks the mean difference between SpO₂ and SaO₂ and the two outer lines mark twice the standard deviation in either direction.

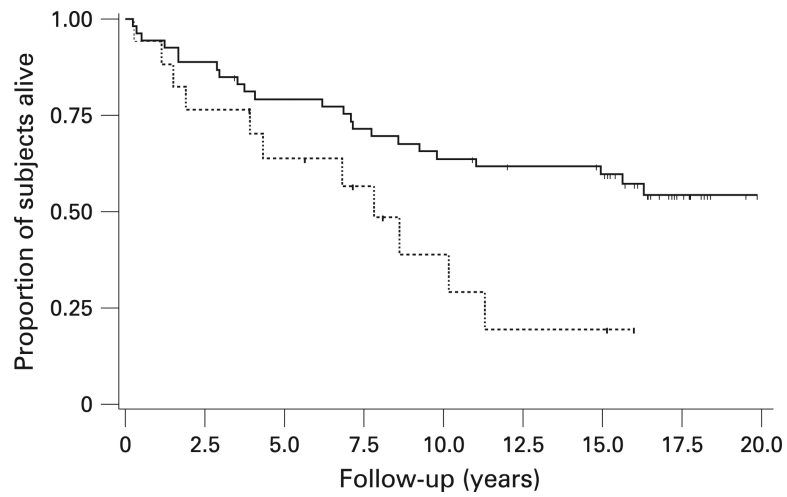


Figure 2. Kaplan–Meier survival curves for subjects with fibrosing interstitial lung disease related to systemic sclerosis (SSc-ILD) stratified on peripheral oxygen saturation at maximal exercise (SpO₂max) <89% or ≥89%. The solid line represents subjects whose SpO₂max remained ≥89% and the dashed line represents subjects whose SpO₂max fell to <89%. Tic = censored observation. Median survival for the entire cohort = 9.5 years.

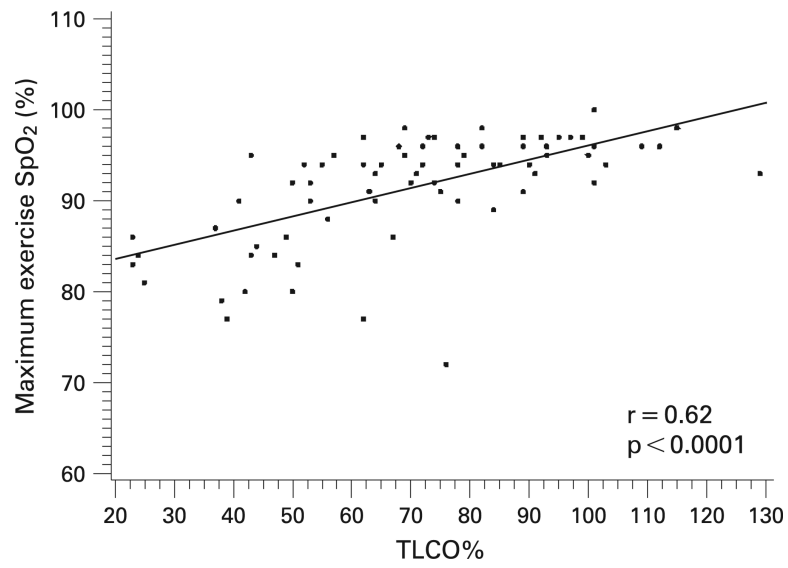


Figure 3. Relationship between peripheral oxygen saturation (SpO₂) at maximum exercise and resting percentage lung carbon monoxide transfer factor (TLCO%).

Table 1

Demographic and clinical characteristics of study subjects

SSc-ILD (n = 83)		
Male	17 (20)	
Mean (SD) age (years)	49.5 (11.5)	
White, n (%)	66 (80)	
Smoking status, n (%)		
Current	9 (11)	
Former	32 (39)	
Never	42 (50)	
Pulmonary physiology		
Rest FVC (%)	70.9 (21.4)	73 (52–89)
Rest TLCO (%)	70.4 (23.7)	72 (41–89)
Surgical biopsy, n (%)	17 (20)	
UIP	6 (35)	
NSIP	4 (24)	
ESHL	4 (24)	
Other	3 (17)	
Exercise variables		
Oxygenation		
Baseline SpO ₂	93.8 (2.6)	94 (93–96)
Baseline Sao ₂	92.3 (2.5)	93 (91–94)
Baseline Pao ₂	71.0 (9.9)	71 (64–78)
Max exercise SpO ₂	91.4 (5.8)	93 (89–96)
Max exercise Sao ₂	89.9 (6.0)	92 (87–94)
Max exercise Pao ₂	70.2 (17.2)	69 (56–83)
Exercise capacity		
$\dot{V}O_2$ max (l/min)	1.0 (0.4)	0.9 (0.7–1.3)
$\dot{V}O_2$ max (%)	56.1 (17.2)	56 (44–65)
Work (Watts)	83.8 (34.2)	75 (60–100)
Work (%)	62.8 (19.2)	62 (38–72)

Data presented as n (%), mean (SD) or median (interquartile range).

FVC, percentage predicted forced vital capacity; ESHL, end stage honeycomb lung; NSIP, non-specific interstitial pneumonia; Pao₂, arterial oxygen tension; SpO₂, peripheral oxygen saturation; Sao₂, arterial blood saturation; SSc-ILD, fibrosing interstitial lung disease related to systemic sclerosis; TLCO, percentage predicted lung carbon monoxide transfer factor; UIP, usual interstitial pneumonia; $\dot{V}O_2$ max, maximum oxygen uptake.